

Université Paris-VI

Biologie génique

Objectifs au cours de

Formation de base IFTAB (2ème année)

Diplôme d'Université P. et M. Curie

Formation continue

2006 - 2007

Pr. A. Raisonnier (raisonni@ccr.jussieu.fr)
Avec l'autorisation des Professeurs J. Etienne et G. Lucotte

Mise à jour : 14 juin 2006
Relecture : Pr. A. Raisonnier


faculté de médecine
Pierre et Marie CURIE

Plan du cours




faculté de médecine
Pierre et Marie CURIE


faculté de médecine
Pierre et Marie CURIE


faculté de médecine
Pierre et Marie CURIE


faculté de médecine
Pierre et Marie CURIE

Objectifs

Des connaissances de base en enzymologie, en virologie et en biochimie des acides nucléiques sont indispensables pour suivre cet enseignement. (Objectifs pré-requis).

1. Enzymes agissant sur les acides nucléiques
    - Pour chacune des enzymes suivantes, utilisées au laboratoire de biologie moléculaire, connaître[1] les facteurs en présence utiles à la réaction (substrat, modèles, produits, co-facteurs, effecteurs, conditions physiques), la spécificité de l'activité enzymatique et donner des exemples de l'usage qu'on peut en faire : polynucléotide kinase, terminal transferase, phosphatase alcaline, DNA polymérases, reverse transcriptase, RNA polymérases, poly(A) polymérase, DNA ligases, RNA ligase, topoisomérase, enzymes de restriction, ribonucléases (A, T1, H), désoxyribonucléases, protéinase K, β-galactosidase.
2. Préparation des acides nucléiques
    - Définir[2] les termes suivants : oligonucléotide, sonde, ADN complémentaire.
    - Décrire les gestes[3] qui permettent de faire l'extraction et la purification d'ADN génomique à partir d'un prélèvement de sang.
    - Expliquer le principe de l'extraction de l'ARN messager d'un tissu et de la réalisation pratique d'une banque de cDNA.
    - Décrire les gestes qui permettent de faire l'amplification par PCR d'un fragment d'ADN dont on possède les amorces.
    - Expliquer le principe de la synthèse d'un oligonucléotide au moyen d'un synthétiseur et de la technique des phosphoramidites.
    - Décrire les gestes qui permettent de faire la synthèse d'une queue de nucléotides à l'extrémité d'un fragment d'ADN.
    - Expliquer le principe du dosage des acides nucléiques. Faire le calcul de la quantité d'acides nucléiques à partir des données brutes fournies par un spectrophotomètre, en tenant compte des dilutions et du volume du milieu considéré.
    - Expliquer le principe de l'hybridation d'une sonde. Faire le calcul de la Tm de cette son-

1. **Connaître**

corps chimique : écrire sa formule développée, énumérer les molécules simples dans une structure complexe et nommer les liaisons qui les unissent, expliquer une expérience mettant en évidence une propriété chimique ou physique ;

image : dessiner un objet ou une structure ;

réaction : écrire l'équation chimique ;

voie métabolique : établir son bilan chimique à partir des réactions de chaque enzyme.

2. **Définir** : préciser dans une phrase concise l'essence d'un objet ou les limites d'un concept en excluant toute notion étrangère et en comprenant toutes les variations possibles de l'objet ou du concept cerné.
3. **Expliquer le principe**, **décrire ou mimer les gestes** ou **faire un schéma explicatif** : préciser le motif de chacun de ces gestes. Etre capable de déceler une erreur qualitative dans la description d'un protocole.

Au niveau des travaux pratiques, les connaissances techniques théoriques étant acquises, on devra être capable de passer à la manipulation sans avoir à apprendre que la partie proprement manuelle.



de à partir de sa séquence et d'une formule de calcul. Etablir les conditions des phases d'un programme de PCR utilisant deux amorces de Tm voisines. Expliquer les étapes des techniques d'hybridation moléculaire qui auront été pratiquées ou expliquées en cours à titre d'exemples : hybridation sur filtre, hybridation in situ, hybridation en solution.

- Décrire le principe technique et les gestes qui permettent de faire le marquage d'une sonde par un atome de ^{32}P ou par un nucléotide marqué.
- Expliquer le principe du séquençage d'un fragment d'ADN au moyen d'un séquenceur et de la technique des *dye primers* ou des *dye terminators* ; décrire et interprèter les résultats[1] apparaissant sur l'écran de l'ordinateur après un séquençage automatique.
- Expliquer les principes de la recherche des éléments cis-régulateurs sur la séquence d'un promoteur (*foot printing*).
- Expliquer le principe des réactions permettant le dosage des ARN transcrits : PCR quantitative, protection à la Rnase.

3. Exemples d'applications

- Décrire et interpréter les résultats d'analyse (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) d'une lésion moléculaire conduisant à une expression anormale du gène considéré : mutation faux-sens, non-sens, décalage du cadre de lecture, protéine tronquée, …
- Expliquer les étapes (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) des techniques ayant permis : un diagnostic génotypique, une analyse génique ou une empreinte génétique : RFLP, carte de restriction, haplotypes, séquençage.

4. Génie génétique

- Définir les termes suivants : vecteur, clonage, polylinker, plasmide.
- Décrire les étapes de l'insertion d'un fragment d'ADN dans un plasmide.
- Expliquer les étapes (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) des techniques ayant permis : une mutagénèse, la transposition d'un gène dans une bactérie, la création d'un animal ou d'une plante transgénique.
- Expliquer le principe de la construction d'un vecteur : *cet objectif servira à tester la faculté pour l'étudiant(e) de faire une synthèse des connaissances de l'ensemble des chapitres précedents.*

1. **Interpreter les résultats** : à partir des données affichées par la machine, après la validation technique de la séquence, savoir décrire la structure de l'acide nucléique séquençé et reconnaître sur cette séquence des lésions moléculaires dont l'interprétation ne nécessite la levée d'aucune ambiguïté.



Partie I

Enzymes agissant sur les acides nucléiques

Rappel des objectifs

- Pour chacune des enzymes suivantes, utilisées au laboratoire de biologie moléculaire, connaître[1] les facteurs en présence utiles à la réaction (substrat, modèles, produits, cofacteurs, effecteurs, conditions physiques), la spécificité de l'activité enzymatique et donner des exemples de l'usage qu'on peut en faire : polynucléotide kinase, terminal transferase, phosphatase alcaline, DNA polymérases, reverse transcriptase, RNA polymérases, poly(A) polymérase, DNA ligases, RNA ligase, topoisomérase, enzymes de restriction, ribonucléases (A, T1, H), désoxyribonucléases, protéinase K, β-galactosidase.

1. **Connaître**

corps chimique : écrire sa formule développée, énumérer les molécules simples dans une structure complexe et nommer les liaisons qui les unissent, expliquer une expérience mettant en évidence une propriété chimique ou physique ;

image : dessiner un objet ou une structure ;

réaction : écrire l'équation chimique ;

voie métabolique : établir son bilan chimique à partir des réactions de chaque enzyme.



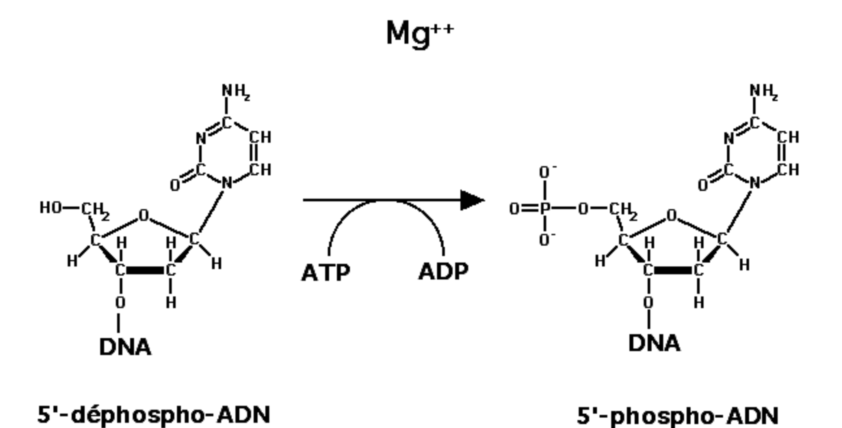

faculté de médecine
Pierre et Marie CURIE

Chapitre 1

Phosphorylation - déphosphorylation



1.1 Polynucléotide-kinase

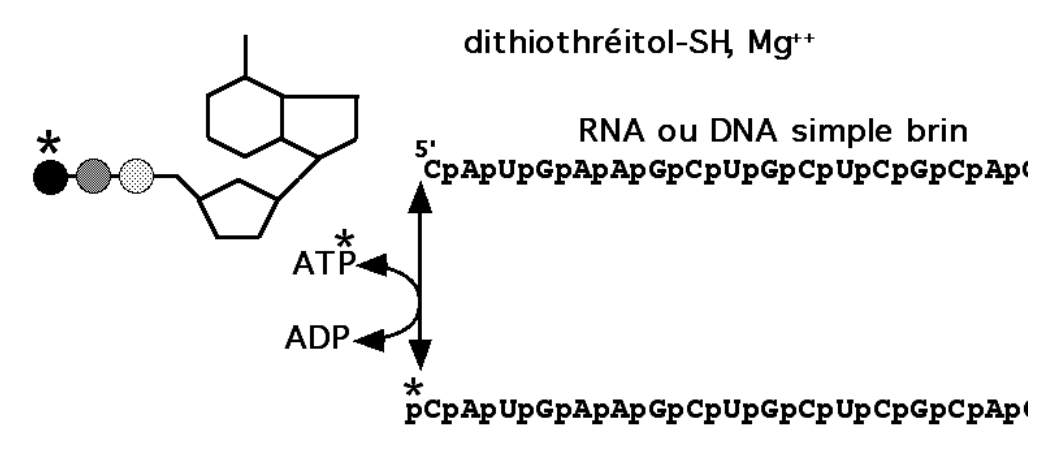


BG 01

- La polynucléotide kinase catalyse le transfert du phosphate γ de l'ATP sur une fonction alcool du carbone 5' d'un ADN ou d'un ARN. Le substrat requiert une déphosphorylation préalable si la fonction est estérifiée.
- L'enzyme catalyse aussi l'échange du phosphate 5' terminal d'un ADN ou d'un ARN avec le phosphate γ de l'ATP, en présence d'ADP comme accepteur de phosphate.
- La polynucléotide kinase du commerce est extraite d'une bactérie (*E. coli*) infectée par le phage T4. Une unité d'enzyme catalyse le transfert de 33 picomoles de phosphate par minute à 37 °C.
- La polynucléotide kinase est utilisée au laboratoire pour incorporer du phosphate radioactif (^{32}P) sur l'extrémité 5' d'un acide nucléique, soit par transfert, soit par échange.



1.2 Polynucléotide kinase T4

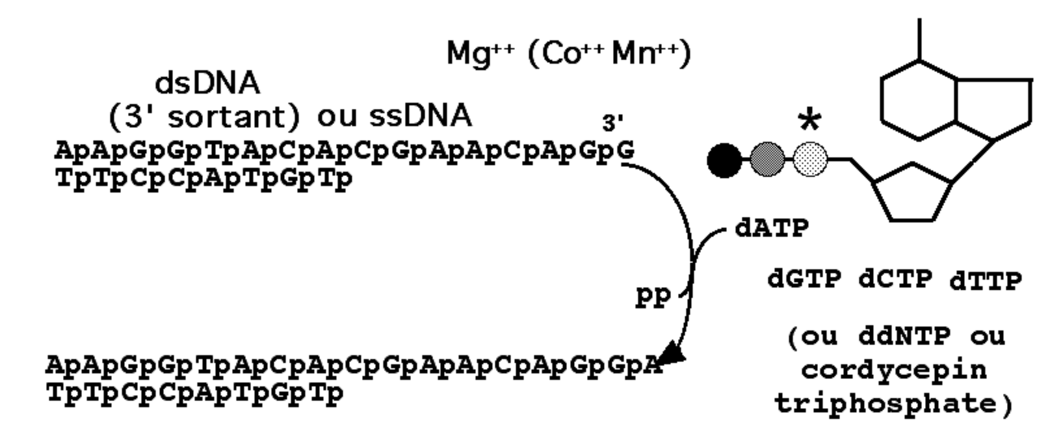


BG 02

- La polynucléotide kinase catalyse le transfert du phosphate γ de l'ATP sur une fonction alcool du carbone 5' d'un ADN ou d'un ARN. Le substrat requiert une déphosphorylation préalable si la fonction est estérifiée. L'enzyme catalyse aussi l'échange du phosphate 5' terminal d'un ADN ou d'un ARN avec le phosphate γ de l'ATP, en présence d'ADP comme accepteur de phosphate.
- La polynucléotide kinase du commerce est extraite d'une bactérie (*E. coli*) infectée par le bactériophage T4. Une unité d'enzyme catalyse le transfert de 33 picomoles de phosphate par minute à 37° C.
- La polynucléotide kinase est utilisée au laboratoire pour incorporer du phosphate radioactif (^{32}P) sur l'extrémité 5' d'un acide nucléique, soit par transfert, soit par échange.



1.3 Terminal transferase



BG 03

- La transférase terminale catalyse l'addition de désoxynucléotides sur une fonction alcool 3'OH terminale de l'ADN. L'enzyme préfère les molécules d'ADN dont l'extrémité 3'OH est sortante, mais il existe des conditions qui favorisent la catalyse sur l'ADN à bouts francs, voire sur les extrémités 3'OH rentrantes. Elle incorpore plus spécifiquement les purines ou les pyrimidines en fonction du cation divalent qu'on lui donne comme cofacteur : Mg^{++} pour les purines, Co^{++} pour les pyrimidines ou encore Mn^{++}.
- La transférase terminale du commerce est extraite du thymus de veau. Une unité d'enzyme catalyse le transfert de 17 picomoles de désoxyribonucléotide par minute à 37° C.
- La transférase terminale est utilisée au laboratoire pour :
  - — construire des séquences polymérisées de nucléotides (queues) à l'extrémité 3'OH de l'ADN (en vue du clonage des fragments d'ADN) ;
  - — marquer les extrémités 3' de l'ADN avec un nucléotide marqué sur le phosphate α ;
  - — ajouter un nucléotide à la fin d'une séquence pour induire une mutation (mutagénèse dirigée).



1.4 Phosphatase alcaline

Intestin de veau **3.1.3.1**

Phosphatase alcaline

Zn^{++}

$$R-PO_3^{2-} \xrightarrow{H_2O} RH + HO-PO_3^{2-}$$

BG 04

- Les phosphatases sont des enzymes à très large spécificité, hydrolysant le radical phosphoryle porté par une liaison ester ou anhydride... On les divise en plusieurs classes en fonction de leur pH optimum d'action : phosphatases acides, phosphatases alcalines. Certaines phosphatases ont une spécificité de substrat plus étroite : 5' nucléotidase. La réaction catalysée par les phosphatases est irréversible, mais certaines sont capables d'échanger le radical phosphate à partir de molécules plus riches en énergie comme le pyrophosphate.
- La phosphatase alcaline du commerce est celle extraite de l'intestin de veau. Une unité d'enzyme catalyse l'hydrolyse d'une micromole de paranitrophénylphosphate par minute à 37 °C.
- La phosphatase alcaline est utilisée au laboratoire pour déphosphoryler les extrémités 5' des acides nucléiques, soit pour préparer un marquage en 5' par une polynucléotide kinase, soit pour empêcher la réannéalisation d'un ADN circulaire linéarisé (vecteurs de clonage).




faculté de médecine
Pierre et Marie CURIE

Chapitre 2

Les polymérases



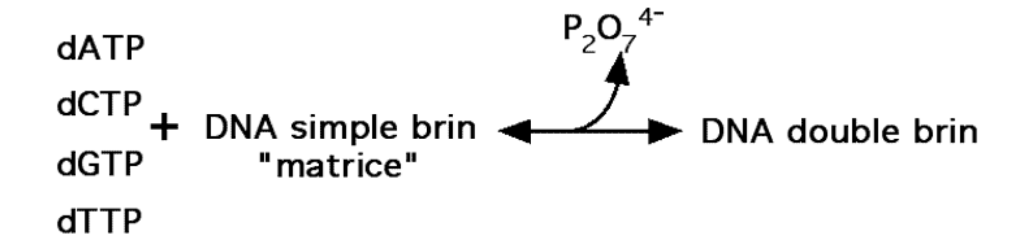
2.1 DNA polymérase (réaction)

2 sous-unités
Isoenzymes

2.7.7.7

DNA polymérase

Primase, DNA-ligase, Hélicase,... Mg^{++}



BG 05

- Les DNA-polymérases sont des enzymes du noyau cellulaire qui agissent en phase S du cycle pour doubler systématiquement l'ensemble du génome diploïde.
- D'autres DNA polymérases interviennent dans la synthèse des amorces RNA/DNA, dans la réplication du génome mitochondrial ou dans la réparation du DNA.
- Elles ne peuvent démarrer la condensation des nucléotides que sur la fonction alcool en 3' du ribose du dernier nucléotide d'une amorce, nucléotide qui doit être hybridé avec le nucléotide complémentaire qui sert de modèle.
- Elles utilisent comme substrats des désoxyribonucléotides triphosphates (dATP, dCTP, dGTP et dTTP) et des amorces de RNA et de DNA synthétisées par une primase (RNA polymérase capable de synthétiser un brin de RNA complémentaire d'un brin de DNA sur 10 nucléotides) suivie d'une DNA polymérase α (qui prolonge l'amorce de RNA de 20 désoxyribonucléotides environ).
- Elles synthétisent l'ADN nouveau par fragments qui, après excision des amorces, sont liés entre eux par une DNA-ligase.
- Elles ont aussi une activité exonucléasique, qui leur permet en particulier d'hydrolyser le dernier nucléotide en 3' du brin synthétisé si celui-ci ne s'apparie pas correctement avec le nucléotide complémentaire du brin modèle.



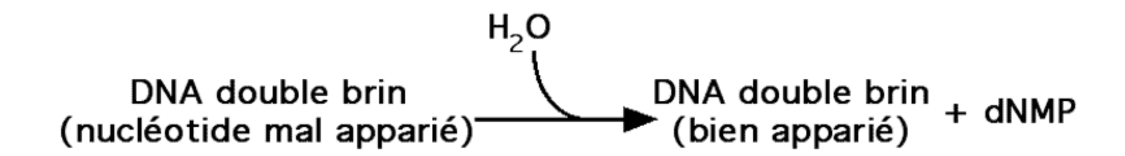
2.2 DNA polymérase (exonucléase 3’→5’)

2 sous-unités
Isoenzymes

3.1.11.1

DNA polymérase
(exonucléase 3'→5')

Primase, DNA-ligase, Hélicase,... Mg^{++}

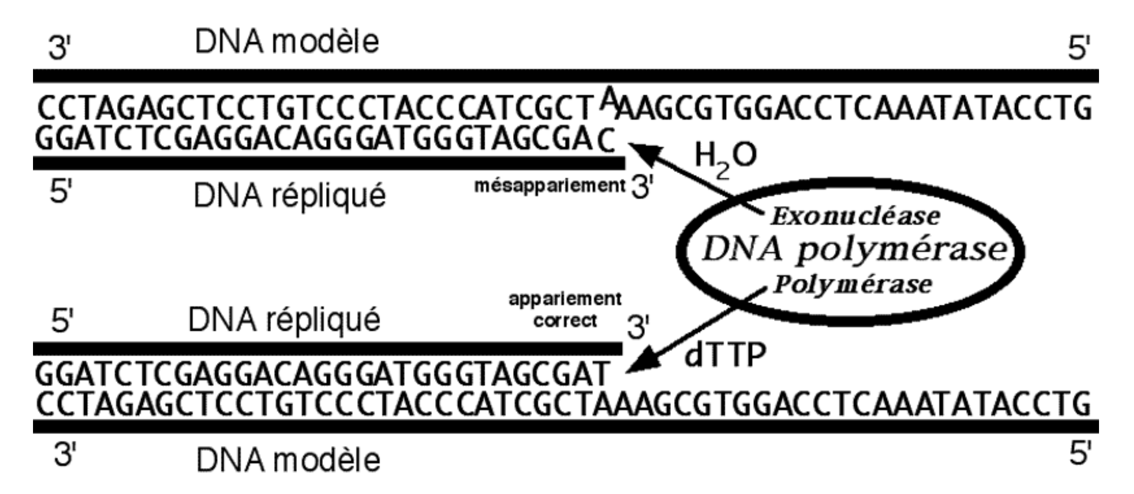


BG 05/1

- Les DNA-polymérases sont des enzymes du noyau cellulaire qui agissent en phase S du cycle pour doubler systématiquement l’ensemble du génôme diploïde.
- D’autres DNA polymérases interviennent dans la synthèse des amorces RNA/DNA, dans la réplication du génome mitochondrial ou dans la réparation du DNA.
- Elles ne peuvent démarrer la condensation des nucléotides que sur la fonction alcool en 3’ du ribose du dernier nucléotide d’une amorce, nucléotide qui doit être hybridé avec le nucléotide complémentaire qui sert de modèle.
- Elles utilisent comme substrats des désoxyribonucléotides triphosphates (dATP, dCTP, dGTP et dTTP) et des amorces de RNA et de DNA synthétisées par une primase (RNA polymérase capable de synthétiser un brin de RNA complémentaire d’un brin de DNA sur 10 nucléotides) suivie d’une DNA polymérase α (qui prolonge l’amorce de RNA de 20 désoxyribonucléotides environ).
- Elles synthétisent l’ADN nouveau par fragments qui, après excision des amorces, sont liés entre eux par une DNA-ligase.
- Elles ont aussi une activité exonucléasique, qui leur permet en particulier d’hydrolyser le dernier nucléotide en 3’ du brin synthétisé si celui-ci ne s’apparie pas correctement avec le nucléotide complémentaire du brin modèle.



2.3 DNA polymérase (fonction d'édition)



BG 05/2

- Les DNA polymérases ont à la fois une activité de polymérase pour ajouter des nucléotides au nouveau brin de DNA, et une activité de 3'→5' exonucléase pour hydrolyser le dernier nucléotide incorporé.
- Si le nucléotide incorporé est complémentaire du nucléotide du brin modèle et donc que l'hybridation des bases se fait normalement, l'activité de polymérase est plus rapide que celle d'exonucléase et le nucléotide suivant va être incorporé.
- Si le nucléotide incorporé n'est pas complémentaire du nucléotide du brin modèle et donc qu'il y a mésappariement, l'activité d'exonucléase est plus rapide que celle de polymérase et ce nucléotide sera hydrolysé.



2.4 DNA-polymérases (tableau)

DNA polymérases	MM	Autres activités 5'→3'	3'→5'	pH opt
E. coli DNA pol I	109000	+	+	7,4
E.coli Klenow fragment	76000	-	+	8,4 (Tris)
T4 DNA pol	114000	-	+	8,5
T7 DNA pol (Sequenase)	92000	-	-	7,7
Th. aquat. DNA pol (Taq)	94000	-	-	8,3 (Tris)
Reverse transcriptase	84000	-	-	7,6

BG 06

- Les DNA polymérases sont des enzymes qui catalysent la synthèse de l'ADN à partir des désoxyribonucléosides triphosphates en recopiant la séquence d'une matrice d'ADN. Certaines d'entre elles ont aussi des activités d'exonucléase qui peuvent s'exercer de 5' vers 3' ou dans l'autre sens, afin de corriger les erreurs d'incorporation (fonction d'édition). Les DNA polymérases sont spécifiques de l'ADN double brin et incorporent les nucléotides pour faire la synthèse d'un deuxième brin en utilisant le premier comme matrice. Les DNA polymérases ont besoin d'une amorce (extrémité 3'OH « rentrante » d'un deuxième brin) pour initier la réaction.
- Tous les êtres vivants ont besoin d'une DNA polymérase pour la réplication de leur ADN. Ce sont en général des protéines d'environ 100000 daltons. Celles des êtres vivants les plus simples (virus, archéobactéries) sont souvent dépourvues de fonctions d'édition. Toutes les DNA polymérases exigent la présence de l'ion magnésium. Beaucoup d'entre elles fonctionnent en milieu alcalin. Leur température optimale d'action se situe habituellement dans une fourchette de 20 à 40 °C.



2.5 DNA pol I (*E. Coli*)

109000
Escherichia coli (lysogène)
2.7.7.7

DNA pol I

Mg++
dsDNA
pGpApGpCpTpG
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp
pTpGpCpCpTpCpTpCpApCpCpCpC
dGTP
dTTP
dATP
dCTP
5'→3' DNA polymerase
pp
5'→3' exonuclease
pT
pG
pC
dCTP
5'→3' DNA polymerase
3'→5' exonuclease
pC
pC
pC
pA
pGpApGpCpTpGpTpApC
pCpTpCpTpC
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp

BG 07

- La DNA polymérase I d'*Escherichia coli* (DNA pol I) est une protéine qui assure la fonction de réplication de l'ADN bactérien.
- Elle initie la réaction à l'extrémité 3' rentrante d'une amorce d'ADN ou d'ARN. Elle incorpore des nucléotides à partir des quatre désoxyribonucléosides triphosphates, en hydrolysant un pyrophosphate et en incorporant le nucléoside et le phosphate α. Lorsque cet atome de phosphore est radioactif, le brin d'ADN produit est aussi marqué.
- DNA pol I est douée d'une double fonction d'édition :
  - — exonucléase 5'→3' pour digérer le deuxième brin à partir de son extrémité 5' (*nick translation*)
  - — exonucléase 3'→5' pour digérer l'extrémité 3' d'un brin (correction immédiate des misappariements)
- Les activités d'exonucléase s'exercent aussi sur les brins d'ARN hybridés (amorces).
- DNA pol I est utilisée pour la synthèse de brins d'ADN marqués sur toute la longueur de la chaîne par la technique de translation de brèche avec des dNTP marqués.



2.6 Fragment de Klenow

76000
produit de protéolyse

2.7.7.7

Klenow fragment

Mg^{++}

dsDNA

pGpApGpCpTpG
pTpGpCpCpTpCpTpCpApCpCpCpC
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp

dGTP
dTTP
dATP
dCTP
5'→3' DNA polymerase
pp

3'→5' exonuclease
dCTP
5'→3' DNA polymerase

pC
pC
pC
pA

GpApGpCpTpGpTpApC
pTpGpCpCpTpCpTpC
CpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp

BG 07/1

- La digestion de la DNA polymérase I par une protéase (subtilisine) donne deux fragments : celui de 76 kD (fragment de Klenow) possède encore deux des activités catalytiques de la polymérase : 5'→3' polymérase et 3'→5' exonucléase. Ce fragment peut-être utilisé pour synthétiser le deuxième brin à partir d'un ADN simple brin et d'une amorce.
- La réaction est la même que celle de la DNA polymérase I. Les conditions de milieu et la vitesse de réaction sont les mêmes que celles de l'enzyme entière.
- Le fragment de Klenow est utilisé pour :

  — la synthèse du deuxième brin, complémentaire d'un cDNA ;
  — le marquage des extrémités 5' sortantes du DNA double brin ;
  — le marquage du DNA par la technique des amorces aléatoires ;
  — le séquençage du DNA par la technique des didésoxynucléotides ;
  — la mutagénèse dirigée à partir d'oligonucléotides synthétiques.



2.7 T4 DNA polymérase

114000
Bacteriophage T4
2.7.7.7
T4 DNA pol
Mg^{++}
dsDNA
pGpApGpCpTpG
pTpGpCpCpTpCpTpCpApCpCpCpC
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp
dGTP
dTTP
dATP
dCTP
5'→3' DNA polymerase
3'→5' exonuclease
pC
pC
pC
dATP
5'→3' DNA polymerase
pp
pGpApGpCpTpGpTpApC
pTpGpCpCpTpCpTpCpA
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTp

BG 08

- Dans une culture d'*E. coli* infectée par le bactériophage T4, on peut purifier la DNA polymérase de ce phage. Comme le fragment de Klenow elle est dépourvue d'activité 5'→3' exonucléase, mais elle possède une activité 3'→5' exonucléase 200 fois plus grande.
- En l'absence de désoxynucléosides triphosphates, seule l'activité 3'→5' exonucléase se manifeste, aboutissant à un DNA avec de longues extrémités 5' sortantes. Lorsque les substrats sont ajoutés au milieu, la resynthèse se fait rapidement, permettant l'incorporation de nucléotides marqués. L'activité de la T4 polymérase est aussi rapide que celle de la DNA polymérase I.
- La DNA polymérase du phage T4 est utilisée pour le marquage du DNA.



2.8 Sequenase

92000
2 sous-unités

2.7.7.7

T7 DNA pol (Sequenase)

Mg++

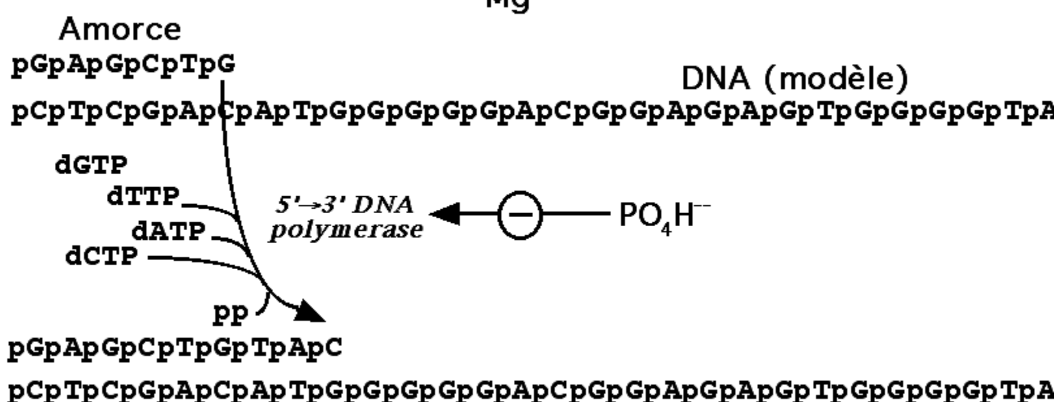
dsDNA
pGpApGpCpTpG
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTpGpGpGpGpTpA
dGTP
dTTP
dATP
dCTP
5'→3' DNA polymerase
ddCTP
3'→5' exonuclease
pp
pGpApGpCpTpGpTpApCpC
pCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTpGpGpGpGpTpA

BG 09

- Les séquenases sont une famille d'enzymes issues de la DNA polymérase du bactériophage T7. Elles sont constituées d'une protéine du phage (gène 5) et d'une protéine de la cellule-hôte (thioredoxine). Elles sont dépourvues par une modification du gène de toute activité d'édition (5'→3' exonucléase et 3'→5' exonucléase).
- Les séquenases sont les plus rapides de toutes les DNA polymérases.
- Les séquenases sont utilisées dans les techniques de séquençage avec des didésoxyribonucléotides (Sanger). Elles sont responsables de quelques erreurs à défaut d'activité d'édition : de l'ordre de 1 misappariement pour 1000 paires de bases.



2.9 Taq polymérase

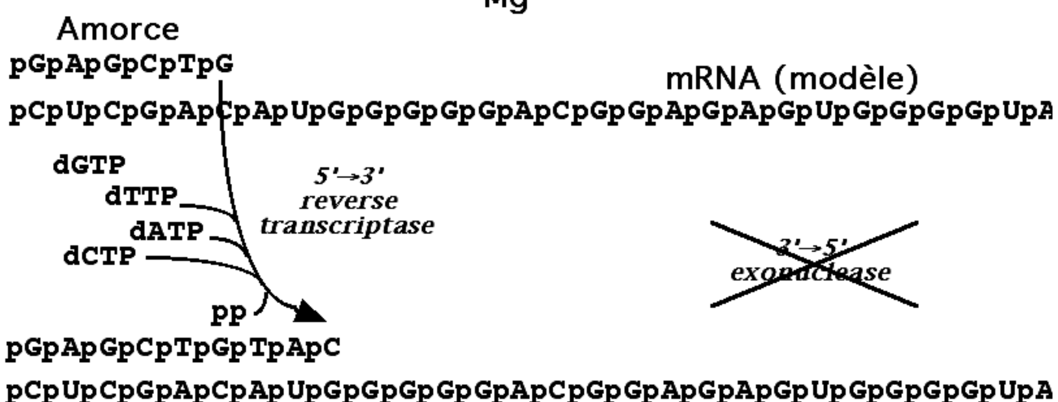


BG 10

- *Thermus aquaticus* est une bactérie thermophile des sources chaudes du parc de Yellowstone (Californie). Elle possède des enzymes thermorésistantes, dont une DNA polymérase (Taq polymérase) résistante à l'ébullition et active à 75-80 °C. La Taq polymérase est dépourvue d'activités d'édition (5'→3' exonucléase et 3' 5' exonucléase).
- La Taq polymérase est inhibée par les ions phosphates.
- La Taq polymérase est utilisée pour la réaction de polymérisation en chaîne (*Polymerase Chain Reaction* = PCR), technique courante d'amplification des fragments de DNA. Parce qu'elle est dépourvue d'activités d'édition la Taq polymérase est responsable de nombreux misappariements : de l'ordre de 1 pour 100 paires de bases.



2.10 Reverse transcriptase



BG 11

- Les transcriptases réverses sont des DNA polymérases qui peuvent synthétiser un brin d'ADN complémentaire (ADNc ou cDNA) en prenant un ARN comme matrice, pour former un hybride ADN:ARN. Elles catalysent donc la réaction inverse de la transcription, d'où le nom de transcriptases réverses.
- Les transcriptases réverses sont produites par des cellules infectées par des rétrovirus, virus à ARN qui font synthétiser un ADNc par la cellule-hôte afin de permettre leur réplication.
- La transcriptase réverse est utilisée pour l'étude des ARN. Après la synthèse de l'ADN complémentaire, on détruit l'ARN matrice, puis on soumet l'ADNc à l'amplification par la PCR, qui fournit une grande quantité d'ADNc hybridé avec une copie ADN de la séquence de l'ARN de départ.



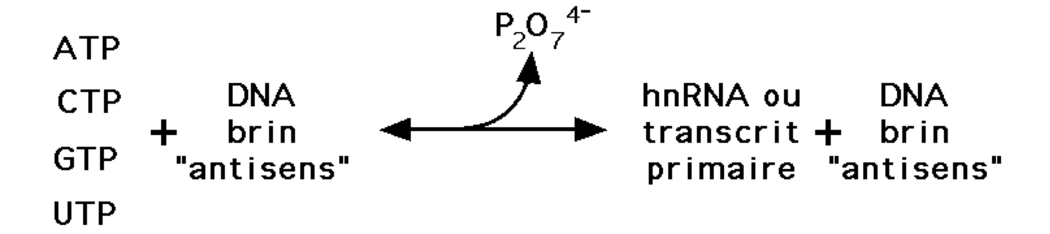
2.11 RNA polymérase II (réaction)

500000
10 sous-unités

2.7.7.6

RNA polymérase II

Facteurs TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, TFIIJ, Mg^{++}



BG 12

- La RNA polymérase II est l'enzyme de la transcription des gènes exprimés sous forme de protéines. Elle est inhibée spécifiquement par un poison extrait de l'Amanite phalloïde, l'α-amanitine.
- Elle est présente dans tous les noyaux cellulaires. Sa masse moléculaire est de 500000 daltons pour 10 sous-unités.
- La transcription qu'elle catalyse nécessite des ribonucléosides triphosphates comme substrats (ATP, CTP, GTP et UTP), plusieurs cofacteurs protéiniques (*transcription factors* TFIIA à TFIIJ). L'énergie de la réaction est fournie par l'hydrolyse des liaisons riches en énergie des nucléosides triphosphates (42 kJ/mol).

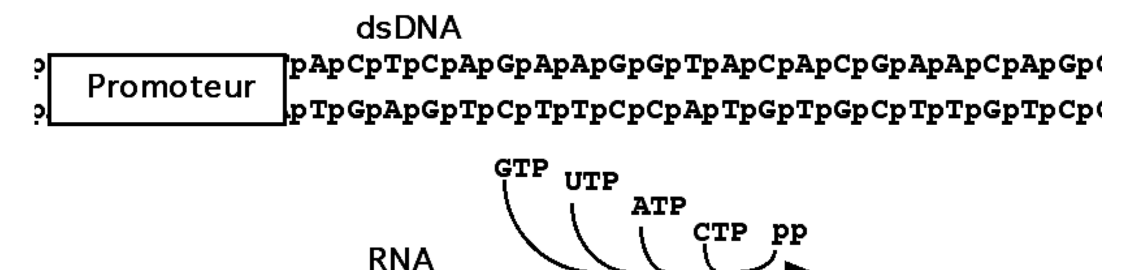


2.12 RNA polymérases (phages)

Bactériophages SP6, T3 et T7 2.7.7.6

RNA polymérases (phages)

Mg^{++}

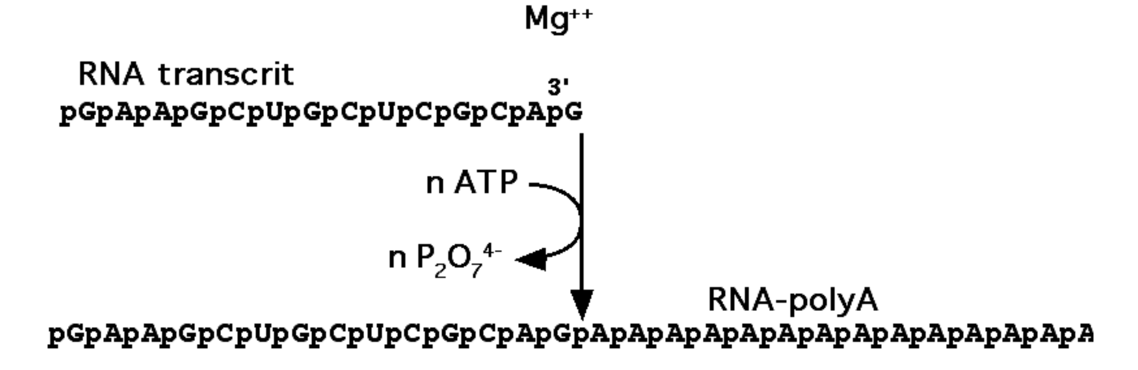


BG 13

- Les RNA polymérases sont les enzymes de la transcription. Elles sont essentielles au cycle des bactériophages à ADN comme le SP6 de *Salmonella typhimurium* ou le phage T7 d'*Escherichia coli*. Elles sont très spécifiques des promoteurs correspondants. Lorsque ces promoteurs sont régulés par un inducteur spécifique, la présence de cet inducteur est indispensable à la réaction.
- Les RNA polymérases des phages sont capables d'incorporer 17 picomoles de nucléotides à la minute à 37 °C.
- Les RNA polymérases des phages sont utilisées pour la transcription in vitro, pour la synthèse des sondes d'ARN et pour l'analyse des messagers (protection à la RNase).



2.13 Poly(A) polymérase



BG 14

- L’extrémité 3’OH terminale du dernier exon du transcrit est d’abord coupée par une nucléase environ 15 nucléotides après la boîte de polyadénylation.
- Ensuite, une polymérase additionne des nucléotides à adénine (plusieurs centaines) pour constituer une « queue » polyA. La réaction se fait sans matrice, à partir de l’ATP et en présence de Magnésium. Le pyrophosphate produit est hydrolysé.
- La queue polyA est indispensable à la sortie du mRNA du noyau vers le cytoplasme. Elle protège le mRNA au cours de la traduction. Les mRNA désadénylés ne sont plus traduits et seront rapidement digérés par la ribonucléase.
- Ces deux activités (nucléase et polymérase) seraient catalysées par un même complexe multienzymatique : polyA polymérase ou polyadénylyl synthétase.

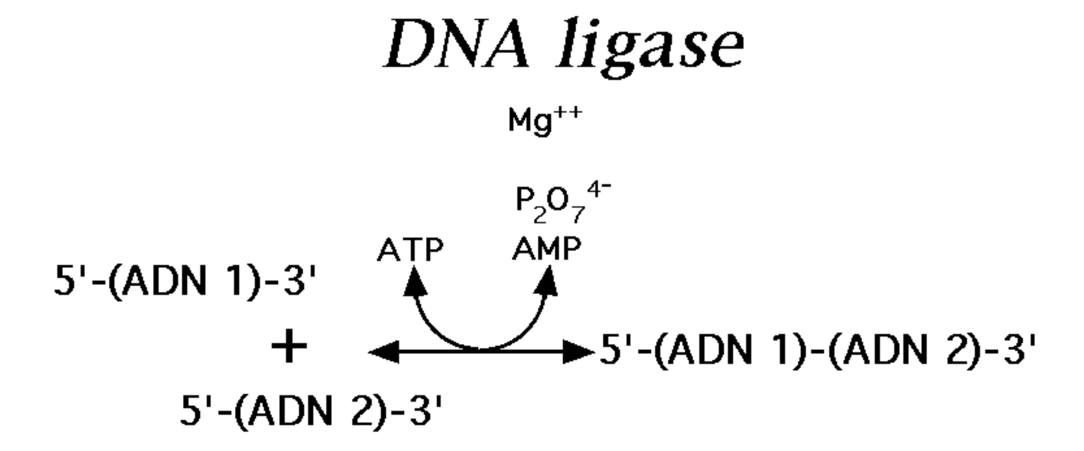


Chapitre 3

Les ligases



3.1 DNA ligase (réaction)



BG 15

- Les DNA ligases sont des enzymes qui sont capables de reconstituer la liaison phosphoester entre le carbone 3’-OH et le phosphate-5’ de deux nucléotides voisins sur un brin de DNA.



3.2 DNA ligase (*E. Coli*)

Escherichia coli *6.5.1.2*

DNA ligase

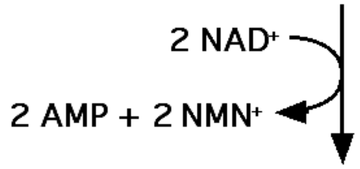
Mg++

dsDNA

DNA, brèche ou bouts collants

GpApGpCpTpGpTpApCpCpCpCpCpT pGpCpCpTpCpTpCpApCpCpC

CpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpA GpApGpTpGpGpG



CpTpApApGpApGpCpTpGpTpApCpCpCpCpCpTpGpCpCpTpCpTpCpApCpCpC

GpApTpTpCpTpCpGpApCpApTpGpGpGpGpGpApCpGpGpApGpApGpTpGpGpG

BG 16

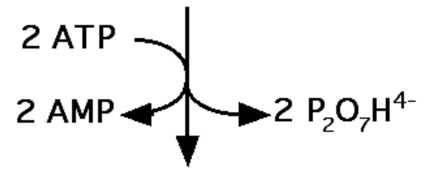
- Les DNA ligases catalysent la liaison de l’extrémité 5’-phosphate terminale d’un brin d’ADN avec l’extrémité 3’-OH terminale d’un autre brin. La DNA ligase de *E. coli* ne lie les bouts francs de l’ADN qu’en présence de réactifs d’exclusion (polyéthylène glycol ou Ficoll), mais elle est toujours active pour souder les fragments d’ADN double brin à extrémités collantes, ou pour fermer une brèche dans un brin d’ADN dont la resynthèse est achevée.
- La DNA ligase de *E. coli* utilise le NAD+ comme coenzyme donneur d’énergie selon la réaction :

$$NAD^+ \rightarrow NMN^+ + AMP$$

- Elle n’a pas d’activité sur les RNA.
- Les DNA ligases sont utilisées dans le clonage des fragments d’ADN et dans la construction des vecteurs.



3.3 DNA ligase (phage T4)



BG 17

- La DNA ligase du bactériophage T4 est moins spécifique que celle de *E. coli* : elle lie bien les fragments de DNA à bouts francs et fonctionne dans la plupart des tampons utilisés par les enzymes de restriction. La T4 DNA ligase agit aussi mais moins rapidement sur les ARN.
- La DNA ligase du phage T4 utilise l'ATP comme donneur d'énergie.



3.4 RNA ligase (phage T4)

Bactériophage T4 *6.5.1.3*

T4 RNA ligase

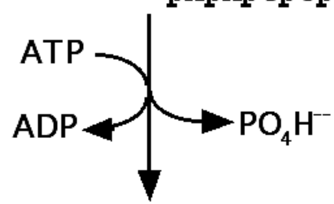
Mg^{++}

RNA ou DNA simple brin

pCpApCpUpGpUpUpApCpCpApApCpApUpG

RNA ou DNA simple brin

pApApGpCpUpGpCpUpCpGpCpApG



pCpApCpUpGpUpUpApCpCpApApCpApUpGpApApGpCpUpGpCpUpCpGpCpAp

BG 18

- La RNA ligase du bactériophage T4 catalyse la liaison des fonctions 5'-phosphate des ADN ou ARN simple brin à la fonction 3'-OH d'autres fragments simple brin d'ARN ou d'ADN. L'enzyme est capable de lier un fragment d'un seul nucléotide.
- La RNA ligase du bactériophage T4 est utilisée pour le marquage des ARN sur leur extrémité 3' et pour la synthèse d'oligonucléotides.



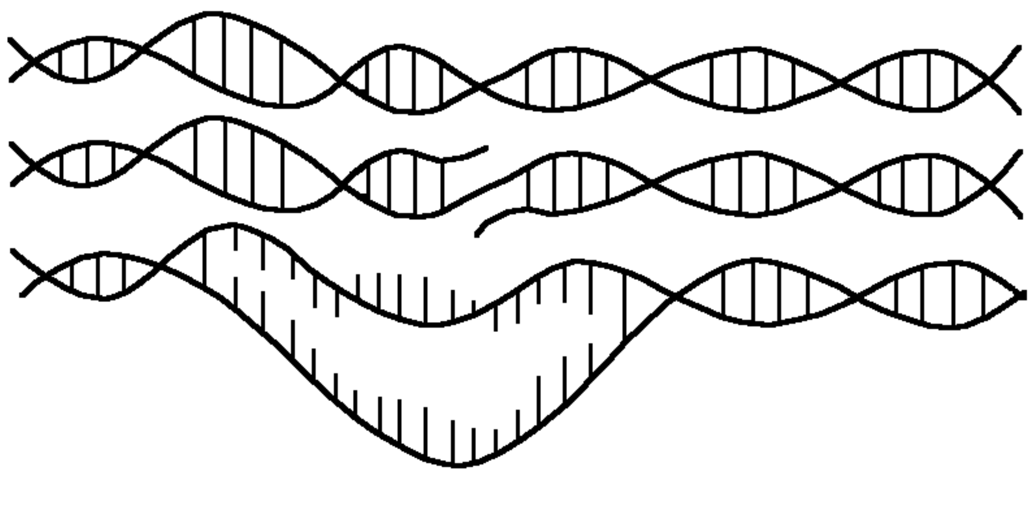

faculté de médecine
Pierre et Marie CURIE

Chapitre 4

Les isomérases



4.1 Topoisomérase



BG 19

- La réplication ou la transcription de l'ADN nécessite une fusion partielle de la double hélice et modifie l'enroulement des deux brins.
- Afin de permettre ces réactions, la topoisomérase est capable de modifier l'enroulement en hydrolysant un brin de l'ADN et en le reconstituant après avoir fait le tour de l'autre brin.
- Après cette opération, la torsion de la double hélice tend à se rapprocher de la valeur normale : pas de l'hélice = 10 paires de nucléotides par tour. Au cours de la réplication et de la traduction le pas de l'hélice diminue en avant des polymérases et l'hélicase (une des topoisomérases) travaille alors pour augmenter le pas. Au contraire en arrière des polymérases les topoisomérases ajoutent des tours pour reconstituer la double hélice.



Chapitre 5

Les nucléases



5.1 Fragment de restriction (définition)

FRAGMENT DE RESTRICTION

- **Produit de la digestion de l'acide désoxyribonucléique par une endonucléase spécifique d'une séquence de nucléotides.**

BG 20

- Les endonucléases de restriction sont des enzymes bactériennes participant à un mécanisme de défense des bactéries vis-à-vis des virus : système de restriction-méthylation.
- Elles catalysent la coupure de l'ADN non méthylé en des endroits caractérisés par une séquence spécifique de nucléotides (site de restriction). Les produits de cette digestion sont les fragments de restriction, dont la longueur, toujours la même pour un ADN donné, ne dépend que de la séquence primaire de cet ADN.
- L'analyse de la longueur des fragments de restriction, à la recherche de variations individuelles (polymorphismes de longueur des fragments de restriction = RFLP) est une des techniques d'analyse de la séquence primaire de l'ADN à la recherche de substitutions, d'insertions ou de délétions qui modifient le nombre de sites de restriction et donc la longueur des fragments de restriction.
- Les extrémités des fragments de restriction peuvent être formées de deux brins d'égale longueur (bouts francs) ou bien présenter un brin plus long que l'autre de quelques nucléotides (bouts collants).



5.2 Système de restriction-modification

RESTRICTION-MODIFICATION :

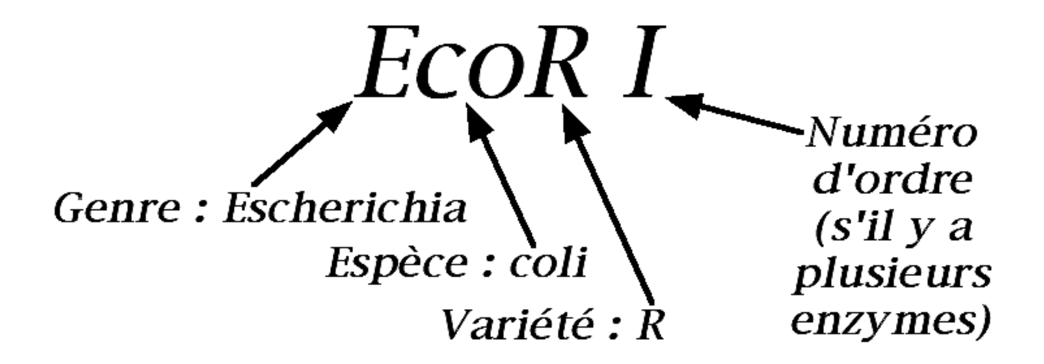
- **Système de défense des Bactéries vis-à-vis des bactériophages, incluant des enzymes de restriction pour digérer l'ADN parasite et des enzymes de méthylation pour protéger l'ADN de la bactérie.**

BG 20/1

- Les Bactéries sont lysées sous l'effet de virus bactériophages dont l'ADN est répliqué par la Bactérie elle-même avant sa destruction.
- Pour détruire l'ADN du parasite la Bactérie exprime des gènes de restriction et de méthylation. Les gènes de restriction permettent la synthèse d'endonucléases coupant l'ADN en des sites très spécifiques.
- Afin de protéger l'ADN bactérien de l'hydrolyse par l'enzyme, une méthylase, codée par le gène de méthylation, va modifier les nucléotides de l'ADN bactérien en les méthylant pour qu'ils ne soient plus reconnus par l'enzyme de restriction.
- L'ensemble du gène de restriction et du gène de méthylation constitue un système de défense de la Bactérie vis-à-vis des phages.

5.3 Nomenclature des enzymes de restriction

3 = Hydrolase **3.1.21.4**
3.1 = Esterase
3.1.21 = Endonucléase produisant un 5'-phosphate
3.1.21.4 = Enzyme avec Mg^{++} (seul cofacteur)



BG 21

- Les enzymes de restriction sont des hydrolases (classe 3 de la E.C.) agissant sur des liaisons esters (sous-classe 3.1), c'est-à-dire des estérases.
- Parmi les estérases on distingue celles qui hydrolysent un acide nucléique en fragments poly-nucléotidiques (endonucléases) et en particulier celles dont les produits gardent leur phosphate 5' initial (sous-sous-classe 3.1.21).
- Enfin en fonction des cofacteurs, les enzymes de restriction ne nécessitent que la présence de l'ion Mg^{++} dans le milieu (3.1.21.4)
- Le nom de chaque enzyme est dérivé du nom d'espèce ou de variété de la Bactérie qui la produit. On écrit l'initiale du nom du genre, les deux initiales du nom de l'espèce, 1 lettre ou 1 nombre pour désigner la variété (ou souche) et après un espace un chiffre romain pour désigner successivement les différentes enzymes de restriction obtenues à partir de la même souche.



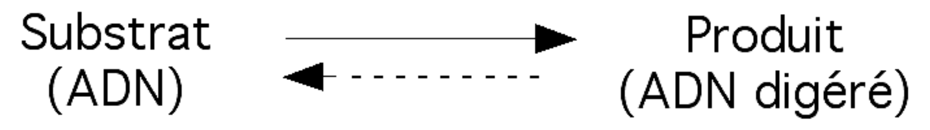
5.4 Restriction (réaction générale)

3.1.21.n

Enzyme (de restriction)

Cofacteurs (Mg^{++}) et coenzymes liés

Conditions physiques : pH (tampon),
potentiel rédox (thiols), force ionique (ions, protéines)



Petites molécules (H_2O)
et Coenzymes libres (S-AdoMet)
Energie

BG 22

- La réaction générale catalysée par les enzymes de restriction implique la présence dans le milieu réactionnel des facteurs suivants :
  - — Enzyme (3.1.21.n)
  - — Substrat : ADN double brin non digéré
  - — Produit : ADN double brin digéré
  - — Cofacteurs
- En général, seul le Mg^{++} est indispensable. Quelques endonucléases font appel à d’autres cofacteurs. Les enzymes de méthylation ont pour coenzyme la S-Adénosyl-Méthionine (S-AdoMet).
- Des facteurs physiques contrôlent aussi ces réactions : pH, potentiel d’oxydoréduction, force ionique. L’énergie libre dégagée par l’hydrolyse est suffisamment importante pour que la réaction ne soit pratiquement pas réversible dans les conditions habituelles.



5.5 Tampons d'incubation (restriction)

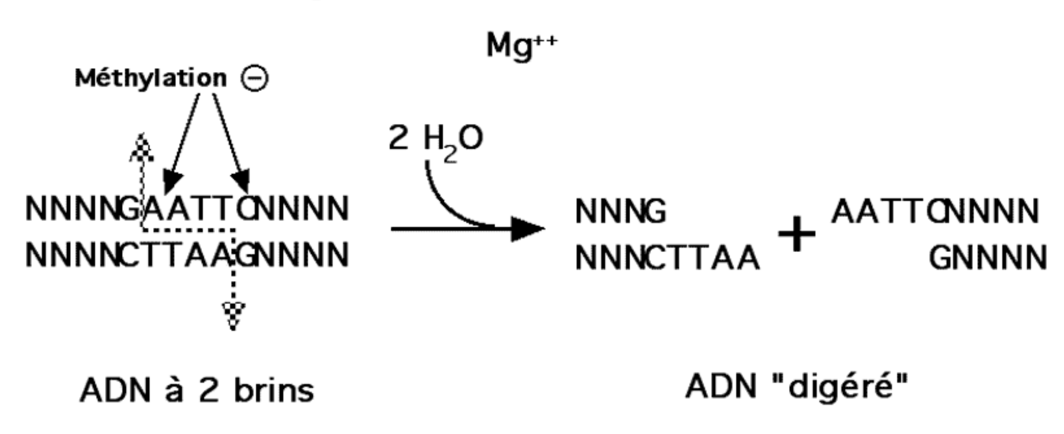
10 mM Tris Propane-HCl (pH 7,0 à 25°C) 10 mM $MgCl_2$ (1 mM DTT)	10 mM Tris-HCl (pH 7,9 à 25°C) 10 mM $MgCl_2$ 50 mM NaCl (1 mM DTT)
50 mM Tris-HCl (pH 7,9 à 25°C) 10 mM $MgCl_2$ 100 mM NaCl (1 mM DTT)	20 mM Tris-CH_3COOH (pH 7,9 à 25°C) 10 mM $Mg(CH_3COO)_2$ 50 mM CH_3COOK (1 mM DTT)

BG 22/1

- Différents tampons sont utilisés pour les enzymes de restriction permettant l'optimisation des réactions de multiples enzymes avec le même tampon.
- Toutes les enzymes de restriction nécessitent la présence de l'ion Mg^{++}
- Le pH (7,0 - 7,9), le potentiel d'oxydo-réduction (dithiothréitol 1 mM), la force ionique (NaCl ou CH_3COOK de 0 à 100 mM) varient d'un tampon à l'autre.
- De rares enzymes nécessitent des conditions spéciales précisées par leur mode d'emploi.



5.6 EcoR I



BG 23

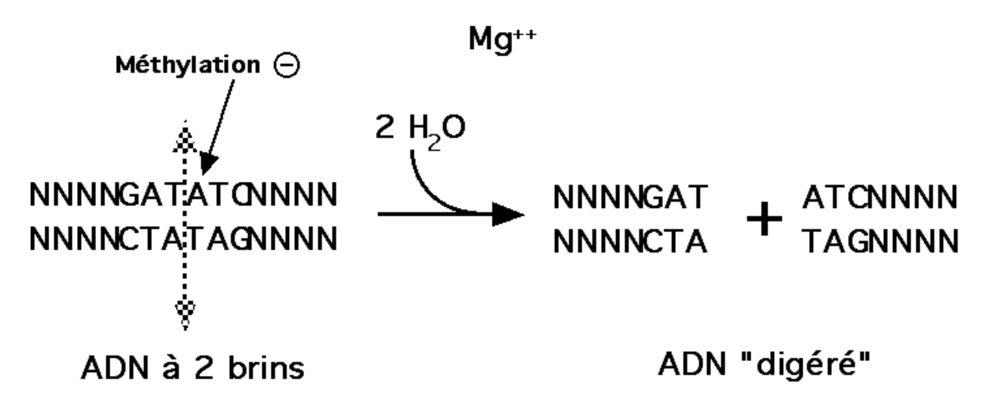
- EcoR I est une enzyme de restriction produite par *Escherichia coli*, souche R.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **GAATTC** 3'
3' **CTTAAG** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait loin du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation des adénines ou de la cytosine du site d'hydrolyse inhibent la reconnaissance du site par EcoR I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.7 EcoR V



BG 24

- EcoR V est une enzyme de restriction produite par *Escherichia coli*, souche R.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **GATATC** 3'

3' **CTATAG** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait au centre de symétrie du palindrome toutes les paires de nucléotides restent appariées : les fragments qui en résultent sont dits « à bouts francs ».
- La méthylation des adénines du site d'hydrolyse inhibe la reconnaissance du site par EcoR V. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.8 Pvu II

3.1.21.4

Pvu II
(enzyme de restriction)

Mg^{++}

Méthylation ⊖

$2\ H_2O$

NNNNCAGCTGNNNN
NNNNGTCGACNNNN

→

NNNCAG
NNNGTC
+
CTGNNNN
GACNNNN

ADN à 2 brins

ADN "digéré"

BG 24/1

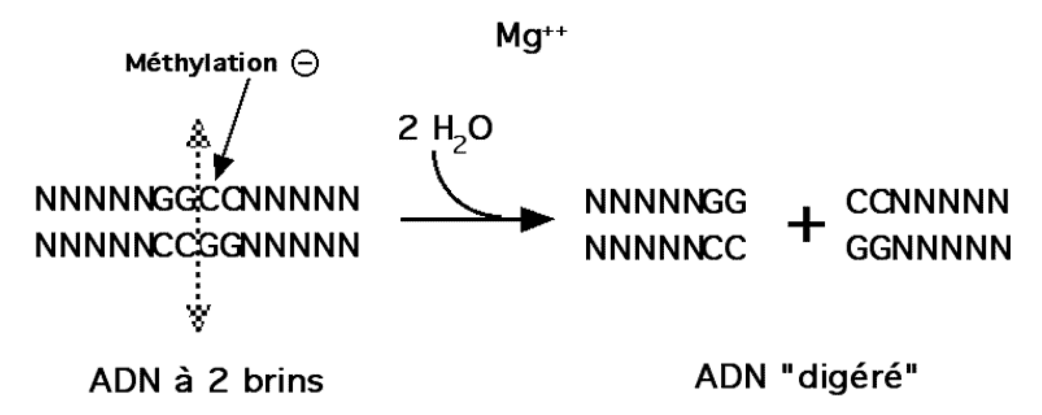
- Pvu II est une enzyme de restriction produite par *Proteus vulgaris*.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **CAGCTG** 3'
3' **GTCGAC** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait au centre de symétrie du palindrome toutes les paires de nucléotides restent appariées : les fragments qui en résultent sont dits « à bouts francs ».
- La méthylation des cytosines au niveau du site d'hydrolyse inhibe la reconnaissance du site par Pvu II. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.9 Hae III



BG 25

- Hae III est une enzyme de restriction produite par *Haemophilus aegypticus*.
- Le site de liaison à l'ADN est formé de quatre paires de nucléotides :

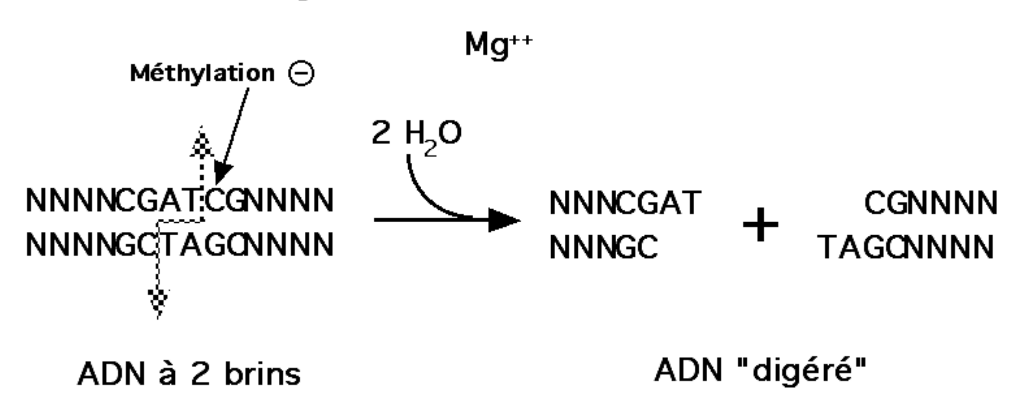
5' GGCC 3'

3' CCGG 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait au centre de symétrie du palindrome toutes les paires de nucléotides restent appariées : les fragments qui en résultent sont dits « à bouts francs ».
- La méthylation de la cytosine immédiatement en aval du site d'hydrolyse inhibe la reconnaissance du site par Hae III. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.10 Pvu I



BG 26

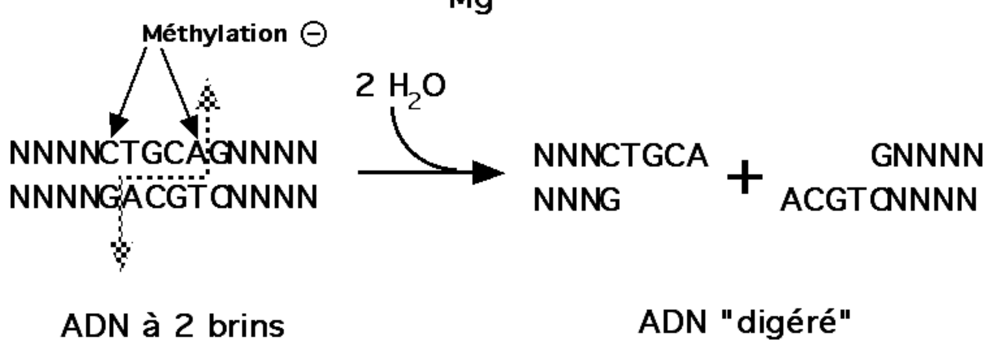
- Pvu I est une enzyme de restriction produite par *Proteus vulgaris*.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **CGATCG** 3'
3' **GCTAGC** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait loin du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation de la deuxième cytosine proche du site d'hydrolyse inhibe la reconnaissance du site par Pvu I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.11 Pst I



BG 26/1

- Pst I est une enzyme de restriction produite par *Providencia stuartii*.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

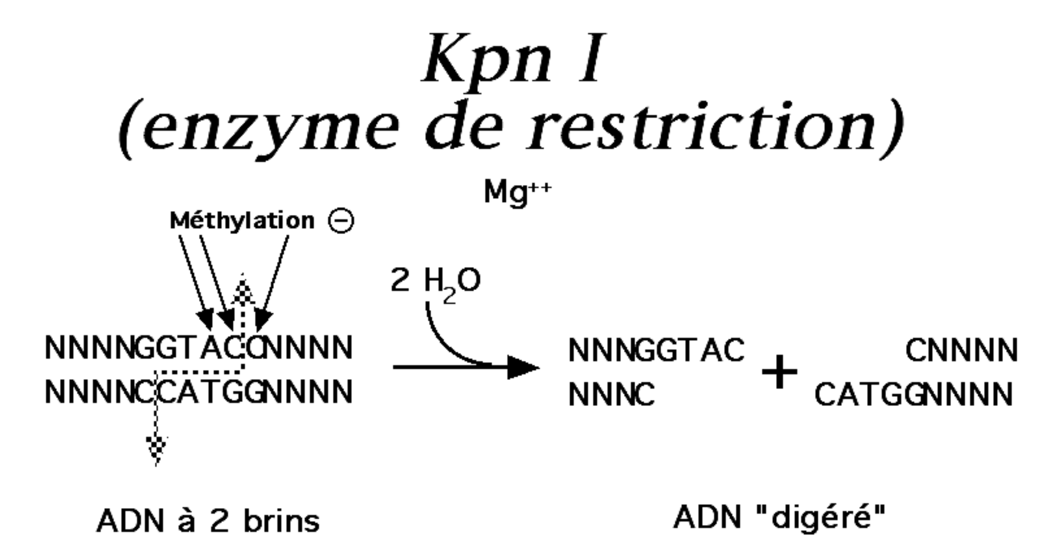
5' **CTGCAG** 3'

3' **GACGTC** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait loin du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation de la première cytosine ou de l'adénine du site d'hydrolyse inhibent la reconnaissance du site par Pst I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.12 Kpn I



BG 26/2

- Kpn I est une enzyme de restriction produite par *Klebsiella pneumoniae*.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **CTGCAG** 3'
3' **GACGTC** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait loin du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation de l'adénine ou des cytosines du site d'hydrolyse inhibent la reconnaissance du site par Kpn I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.13 Alphabet dégénéré

Alphabet dégénéré du code génétique

A = Adénosine
C = Cytidine
G = Guanosine
T = Thymidine

R = A ou G puRine
Y = C ou T pYrimidine

B = non A (C,G ou T)
D = non C (A,G ou T)
H = non G (A,C ou T)
V = non T (A,C ou G)

K = G ou T cétone (Keto en anglais)
M = A ou C aMine
N = A,C,G ou T

S = C ou G 3 liaisons H (Strong en anglais)
W = A ou T 2 liaisons H (Weak en anglais)

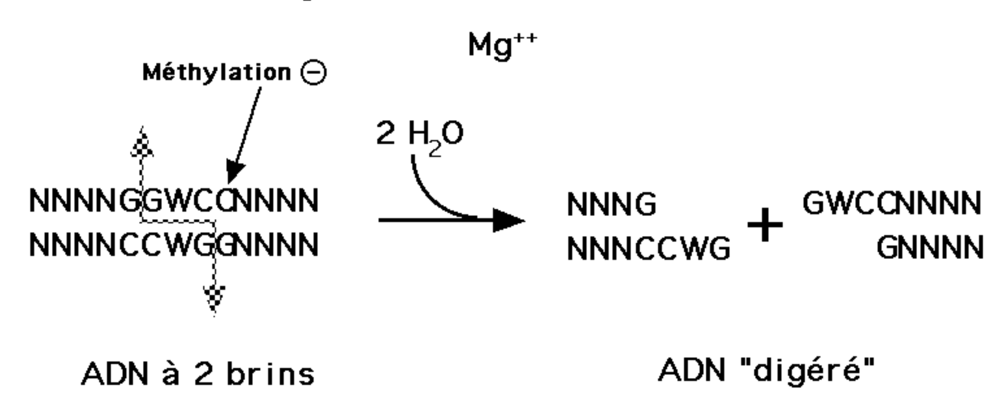
Le complément de :	A	C	G	T	R	Y	K	M	S	W	B	D	H	V	N
est :	T	G	C	A	Y	R	M	K	S	W	V	H	D	B	N

BG 27

- La spécificité large de certaines enzymes ou encore la dégénérescence du code génétique implique qu'on doit indiquer quelquefois dans les séquences des lettres correspondant à plusieurs bases azotées différentes pour le même nucléotide présent à une position.
- Le choix peut porter sur deux des 4 nucléotides : A ou G, C ou T, G ou T, A ou C, C ou G, A ou T ; ou bien sur 3 nucléotides : C, G ou T, A, G ou T, A, C ou T, A, C ou G ; ou enfin sur les 4 nucléotides : A, C, G ou T.
- Les règles de complémentarité impliquent évidemment une dégénérescence de la séquence de l'autre brin en regard d'une position dégénérée.



5.14 Ava II



BG 28

- Ava II est une enzyme de restriction produite par *Anabaena variabilis*.
- Le site de liaison à l'ADN est formé de cinq paires de nucléotides :

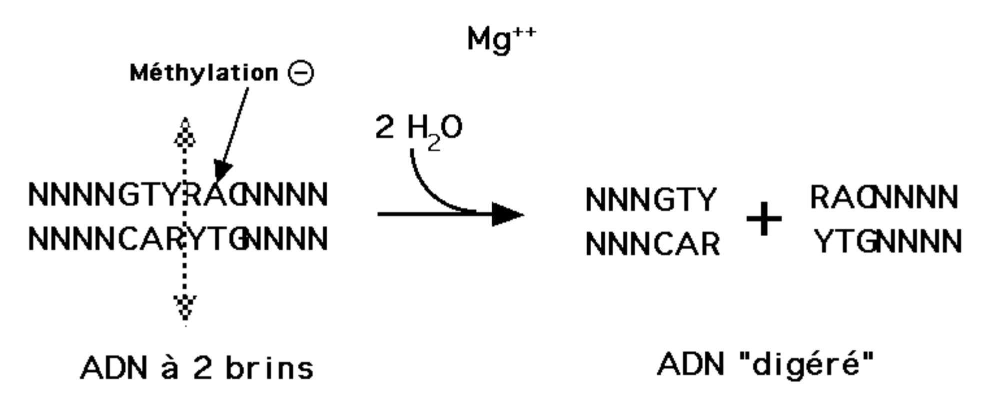
5' **GGWCC** 3'

3' **CCWGG** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait loin du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation des cytosines du site d'hydrolyse inhibe la reconnaissance du site par Ava II. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.15 Hind II



BG 29

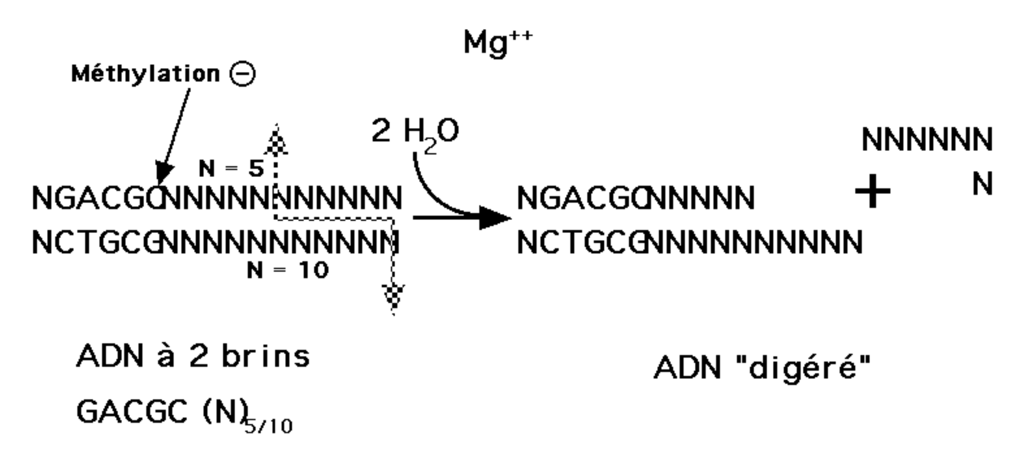
- Hind II est une enzyme de restriction produite par *Haemophilus influenzae*, souche Rd.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **GTYRAC** 3'
3' **CARYTG** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait au centre de symétrie du palindrome toutes les paires de nucléotides restent appariées : les fragments qui en résultent sont dits « à bouts francs ».
- La méthylation de l'adénine du site d'hydrolyse inhibe la reconnaissance du site par Hind II. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.16 Hga I



BG 30

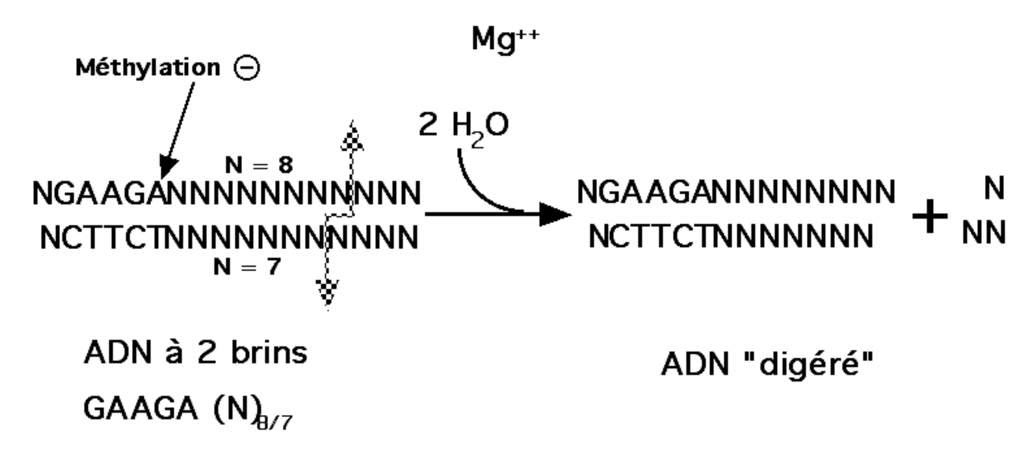
- Hga I est une enzyme de restriction produite par *Haemophilus gallinarum*.
- Le site de liaison à l'ADN est formé de cinq paires de nucléotides :

5' **GACGC** 3'

3' **CTGCG** 5'

- Les sites de restriction ne sont quelquefois pas des palindromes, c'est à dire qu'ils sont différents sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). L'hydrolyse des liaisons phosphodiester se fait en dehors du site de liaison et le site hydrolysé n'est pas dans le site de restriction. Pour indiquer le site de coupure on ajoute à la séquence du site de restriction une fraction indiquant le nombre de nucléotides en aval de la séquence qui séparent l'extrémité 3' de la séquence du site de coupure sur le brin de la séquence et sur le brin complémentaire : ici, 5 nucléotides en aval de la séquence sur le brin de la séquence et 10 nucléotides sur le brin complémentaire. Ici, il va rester des nucléotides non appariés : les fragments de restriction sont donc « à bouts collants ».
- La méthylation de la deuxième cytosine du site d'hydrolyse inhibe la reconnaissance du site par Hga I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.

5.17 Mbo II



BG 31

- Mbo II est une enzyme de restriction produite par *Moraxella bovis*.
- Le site de liaison à l'ADN est formé de cinq paires de nucléotides :

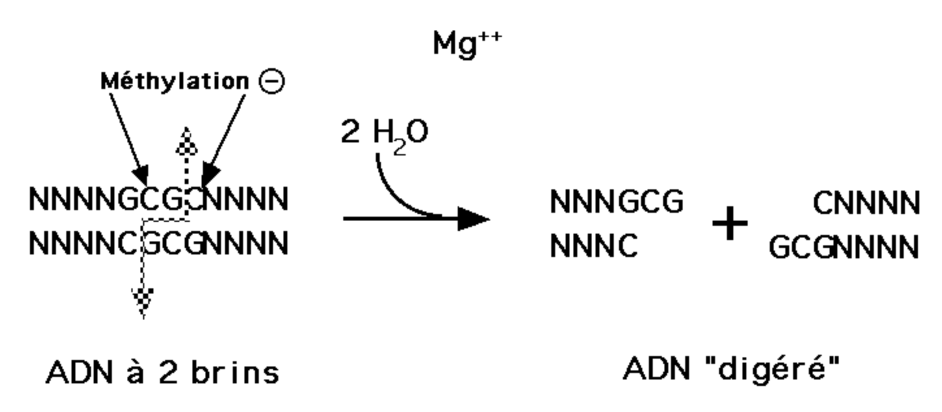
5' **GAAGA** 3'

3' **CTTCT** 5'

- Les sites de restriction ne sont quelquefois pas des palindromes, c'est à dire qu'ils sont différents sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). L'hydrolyse des liaisons phosphodiester se fait en dehors du site de liaison et le site hydrolysé n'est pas dans le site de restriction. Pour indiquer le site de coupure on ajoute à la séquence du site de restriction une fraction indiquant le nombre de nucléotides en aval de la séquence qui séparent l'extrémité 3' de la séquence du site de coupure sur le brin de la séquence et sur le brin complémentaire : ici, 8 nucléotides en aval de la séquence sur le brin de la séquence et 7 nucléotides sur le brin complémentaire.
- La méthylation de la dernière adénine du site d'hydrolyse inhibe la reconnaissance du site par Mbo II. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.18 Hha I



BG 32

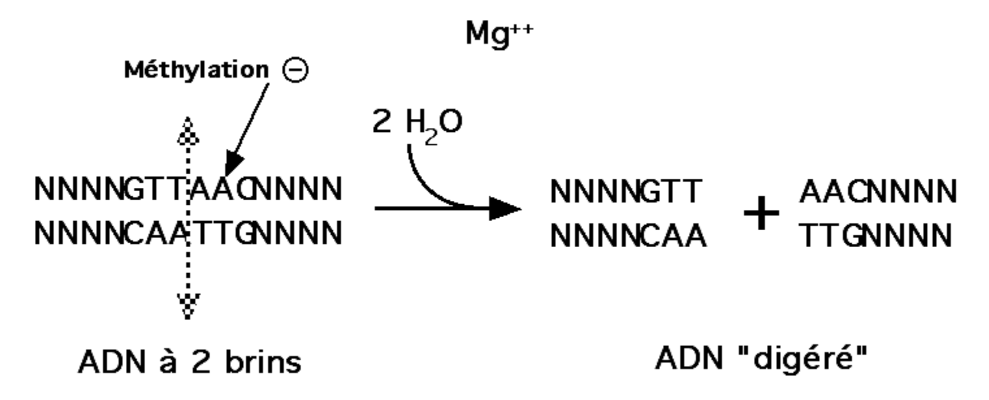
- Hha I est une enzyme de restriction produite par *Haemophilus haemolyticus*.
- Le site de liaison à l’ADN est formé de quatre paires de nucléotides :

5’ **GCGC** 3’
3’ **CGCG** 5’

- Les sites de restriction sont souvent de type palindrome, c’est à dire qu’ils sont identiques sur les deux brins de l’ADN (mais antiparallèles sur l’autre brin). Lorsque l’hydrolyse des liaisons phosphodiester se fait en dehors du centre de symétrie du palindrome certaines paires de nucléotides restent non appariées : les fragments qui en résultent sont dits « à bouts collants ».
- La méthylation des cytosines au niveau du site d’hydrolyse inhibe la reconnaissance du site par Hha I. L’ADN de la bactérie ainsi méthylé n’est pas hydrolysé alors que l’ADN parasite qui n’est pas méthylé sera hydrolysé.



5.19 Hpa I



BG 33

- Hpa I est une enzyme de restriction produite par *Haemophilus parainfluenzae*.
- Le site de liaison à l'ADN est formé de six paires de nucléotides :

5' **GTTAAC** 3'
3' **CAATTG** 5'

- Les sites de restriction sont souvent de type palindrome, c'est à dire qu'ils sont identiques sur les deux brins de l'ADN (mais antiparallèles sur l'autre brin). Lorsque l'hydrolyse des liaisons phosphodiester se fait au centre de symétrie du palindrome toutes les paires de nucléotides restent appariées : les fragments qui en résultent sont dits « à bouts francs ».
- La méthylation de la deuxième adénine de la séquence inhibe la reconnaissance du site par Hpa I. L'ADN de la bactérie ainsi méthylé n'est pas hydrolysé alors que l'ADN parasite qui n'est pas méthylé sera hydrolysé.



5.20 Digestion d’une séquence (Hae III)

```
CTAAGAGCTGTACCCCCTGCCTCTCACCCCATCACCATGAGTCTTCCATGTGCTTG
TCCTCTCCTCCCCCATTTCTCCAACTTGTTTATCCTCACATAATCCCTGCCCCACT= 737 nt
GG
  CCCATCCATAGTCCCTGTCACCTGACAGGGGTGGGTAAACAGACAGGTATATAG
CCCCTTCCTCTCCAGCCAGGGCAGGCACAGACACCAAGGACAGAGACGCTGGCTAG= 152 nt
GTAAGATAAGGAGGCAAGATGTGTGAGCAGCATCCAAAGAGG
                                          CCTGGGCTTCAGTT
GTGGAGAGGGAGAGAGCCAGGTTGGAATGGGCAGCAGGTAGGGAGATCCCTGGGGA= 128 nt
GGAGCTGAAGCCCATTTGGCTTCAGTGTCCCCCAAACCCCCACCACCCTCTTCCTA
GG
  CCGCCCTCCCCACTGTTACCAACATGAAGCTGCTCGCAGCAACTGTGCTACTCC= 98 nt
TCACCATCTGCAGCCTTGAAGGTGGGTGTGATATGGAGAGGGGG
                                            CCAAAAGGGAGA
ATGTGCTGGTGGTTATAGCTACCTATCTGCCTGTGCTCTTATATCCAGGTCTGGCA
ACCAAAGGGCTCAAAAGGAAGATCATTCCCTCTCTAAAGCCCAGAAAGCCCATCCA= 198 nt
AAGATCTGGCTGGATACCTTTTTGGGGAGGGGGCAGAGAGCTGGGGTGGGTACTGG`
CAGAGGTATGGCAGGAGG
                  CCAAGATTCACTGCTGTGGACCCAGCTGAAAAGAGTGT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGGGCAGGAGCTTTGGTTCGGAGACAGGC= 195 nt
AAAGGAGCCATGTGTGGAGAGCCTAGTTTCTCAGTACTTCCAGACCGTGACTGACT
ATGGCAAGGACCTGATGGAGAAGGTCAAGAGCCCAGAGCTTCAGG
                                             CCGAGG     = 6 nt
                                                    CCAA
GTAAGTCTCAGGGCAAGGGGTTCAGGGGCTGTGGAACTGTGGAGAGAAAGAAGGGA= 196 nt
AGATGAGAGGTCCCACAGAAGTCTGAACCCAGGGGTGGGGATTAGGGCAGATTAGG
CTTAAATTGCAGAGAAAAAGTATTTCATCACCCAAAGATCCCACACGTCTCTTAGA
```

BG 34

- Voici le résultat de la digestion du gène de l’apoA-II (brin codant) par Hae III.
- Chacune des séquences GGCC engendre une coupure de l’ADN, ce qui produit des fragments de restriction dont le longueur dépend de la fréquence de cette séquence sur l’ADN.
- On peut établir des cartes des emplacements de ces sites sur l’ADN (cartes de restriction).
- La recherche de ces sites sur l’ADN extrait d’un sujet permet de détecter des séquences différentes qui ajoutent ou suppriment des sites de restriction. Il en résulte que les fragments de restriction chez ces individus n’ont pas la même longueur que les mêmes fragments chez les autres individus : il y a un polymorphisme de longueur des fragments de restriction (RFLP = *Restriction Fragment Length Polymorphism*).



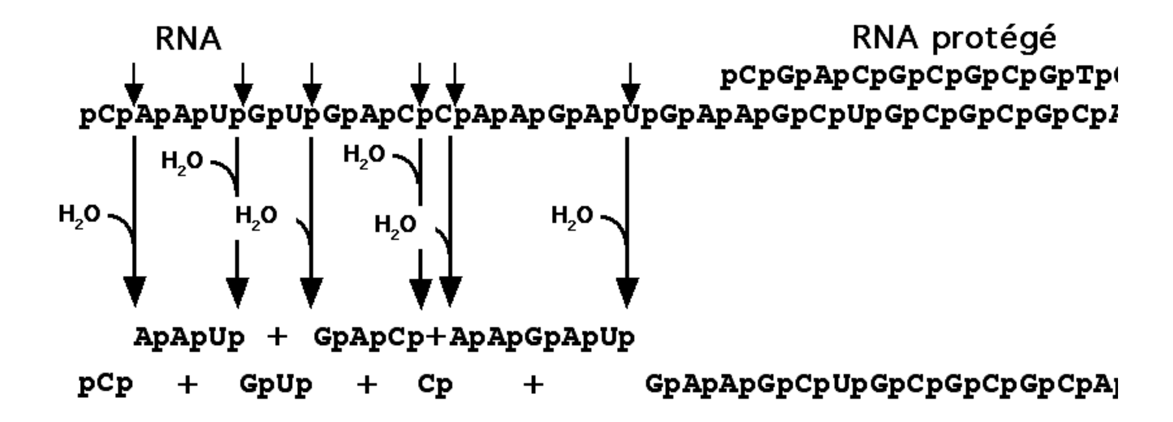
5.21 Digestion d’une séquence (Mbo II)

```
CTAAGAGCTGTACCCCCTGCCTCTCACCCCATCA                            = 299 nt
                                  CCATGAGTCTTCCATGTGCTTGTCCTCTC
CTCCCCCATTTCTCCAACTTGTTTATCCTCACATAATCCCTGCCCCACTGGCCCATCCATAGT
CCCTGTCACCTGACAGGGGTGGGTAAACAGACAGGTATATAGCCCCTTCCTCTCCAGCCAGGG
CAGGCACAGACACCAAGGACAGAGACGCTGGCTAGGTAAGATAAGGAGGCAAGATGTGTGAGC = 343 nt
AGCATCCAAAGAGGCCTGGGCTTCAGTTGTGGAGAGGGAGAGAGCCAGGTTGGAATGGGCAGC
AGGTAGGGAGATCCCTGGGGAGGAGCTGAAGCCCATTTGGCTTCAGTGTCCCCCAAACCCCC
                                                              A
CCACCCTCTTCCTAGGCCGCCCTCCCCACTGTTACCAACATGAAGCTGCTCGCAGCAACTGTG
CTACTCCTCACCATCTGCAGCCTTGAAGGTGGGTGTGATATGGAGAGGGGGCCAAAAGGAGA = 193 nt
ATGTGCTGGTGGTTATAGCTACCTATCTGCCTGTGCTCTTATATCCAGGTCTGGCAACCAAAG
GGCTCAAAAGGAAGATCATTCCC
                       TCTCTAAAGCCCAGAAAGCCCATCCAAAGATCTGGCTGGA
TACCTTTTTGGGGAGGGGGCAGAGAGCTGGGGTGGGTACTGGCAGAGGTATGGCAGGAGGCCA
AGATTCACTGCTGTGGACCCAGCTGAAAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GGGCAGGAGCTTTGGTTCGGAGACAGGCAAAGGAGCCATGTGTGGAGAGCCTAGTTTCTCAGT = 372nt
ACTTCCAGACCGTGACTGACTATGGCAAGGACCTGATGGAGAAGGTCAAGAGCCCAGAGCTTC
AGGCCGAGGCCAAGTAAGTCTCAGGGCAAGGGGTTCAGGGGCTGTGGAACTGTGGAGAGAAAG
AAGGGAAGATGAGAGGT
                 CCCACAGAAGTCTGAACCCAGGGGTGGGGATTAGGGCAGATTAGGC
TTAAATTGCAGAGAAAAAGTATTTCATCACCCAAAGATCCCACACGTCTCTTAGATAGAGAGG
AACAGCAAGAACTGGGCCTTGAATTTCAGTCTCTAGAGTCTGTCCCTCTACCTAGCAAAGGTC
TTGACTCTATTCCTACCTAGGGGCTTTGCCATGCGATGGACCAGGCACTAGAGTTTGGGGACC = 475nt
TGAGTCAGTCTGCTCTGACCTCCCACCACCACCAAGGCCCCTGCCAGTGCCTAGGGTCCCTCA
GATTAAACTCTAATCCCCTCACCTATCCAGGTCTTACTTTGAAAAGTCAAAGGAGCAGCTGAC
```

BG 34/1

- Voici le résultat de la digestion du gène de l’apoA-II (brin codant) par Mbo II.
- Mbo II est spécifique d’un site non palindromique donc dissymétrique : GAAGA $(N)_{8/7}$. Chacune des séquences GAAGA engendre une coupure de l’ADN, à huit nucléotides coté 3’ du site. La séquence doit être recherchée sur les deux brins, ou bien on peut rechercher la séquence complémentaire TCTTC qui engendre une coupure à sept nucléotides coté 5’.
- On peut établir des cartes des emplacements de ces sites sur l’ADN (cartes de restriction).
- La recherche de ces sites sur l’ADN extrait d’un sujet permet de détecter des séquences différentes qui ajoutent ou suppriment des sites de restriction. Il en résulte que les fragments de restriction chez ces individus n’ont pas la même longueur que les mêmes fragments chez les autres individus : il y a un polymorphisme de longueur des fragments de restriction (RFLP = *Restriction Fragment Length Polymorphism*).

5.22 Ribonucléase A

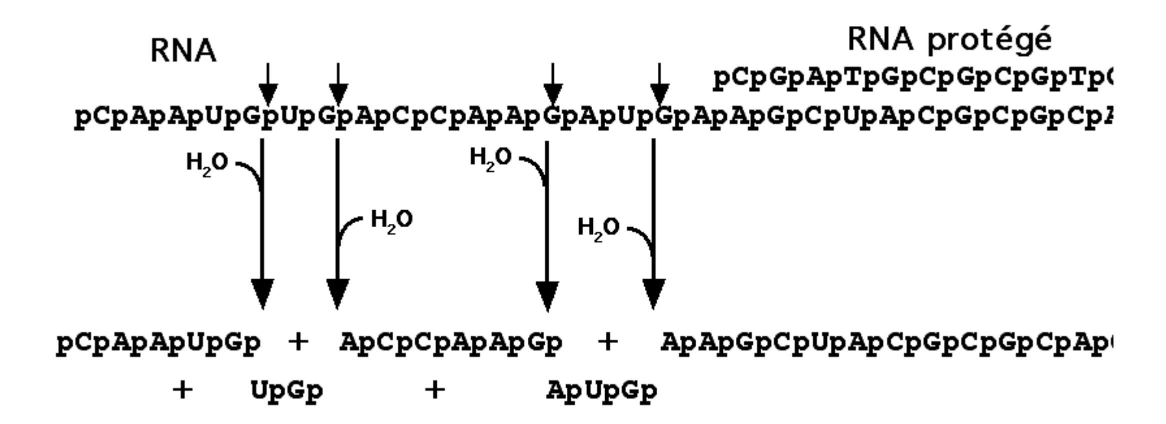


BG 35

- La ribonucléase A est l'enzyme de la digestion des ARN chez les animaux. Elle agit comme une endonucléase, préférentiellement après les nucléotides à pyrimidine, en hydrolysant la liaison entre le phosphate et le carbone 5' du nucléotide suivant. Elle hydrolyse les ARN jusqu'à un mélange d'oligonucléotides se terminant tous par un nucléotide à pyrimidine estérifié par un phosphate en 3'.
- La ribonucléase pancréatique est une des plus petites et des mieux connues des enzymes. Elle est thermorésistante et extrêmement active.
- Il existe des inhibiteurs de la RNase : soit des détergents comme le SDS (laurylsulfate de sodium), soit des protéines comme celle extraite du placenta qui est souvent utilisée pour protéger les ARN dans les réactions enzymatiques.



5.23 Ribonucléase T1



BG 36

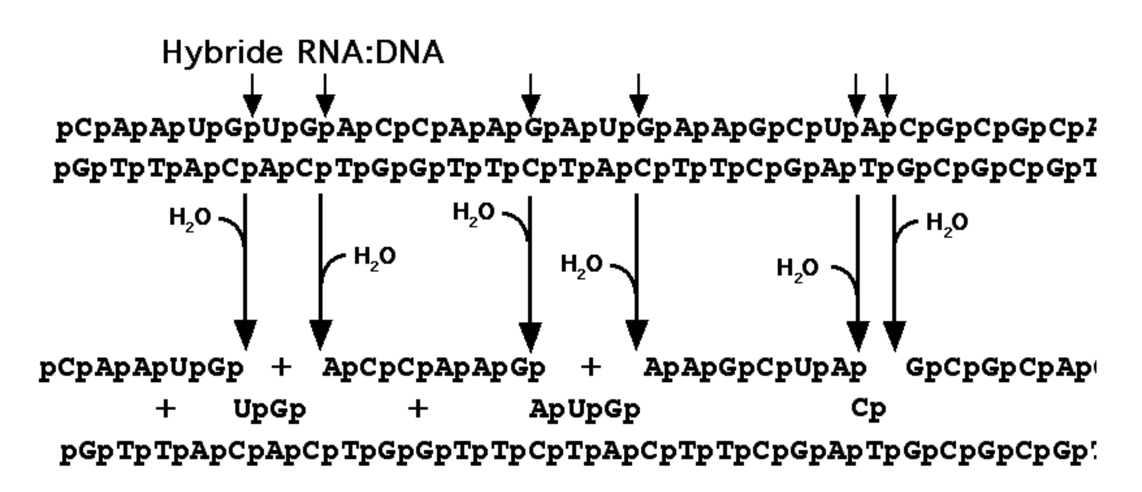
- La ribonucléase T1 hydrolyse spécifiquement les ARN en rompant les liaisons 3'-5' phosphodiester en aval des GMP, de telle manière que le produit soit une guanosine 3'-phosphate ou un oligonucléotide se terminant par une guanosine 3'-phosphate.
- Elle est utilisée pour hydrolyser les ARN non-hybridés lors des expériences d'hybridation ARN:ADN.



5.24 Ribonucléase H

Thymus de veau *3.1.26.4*

Ribonucléase H

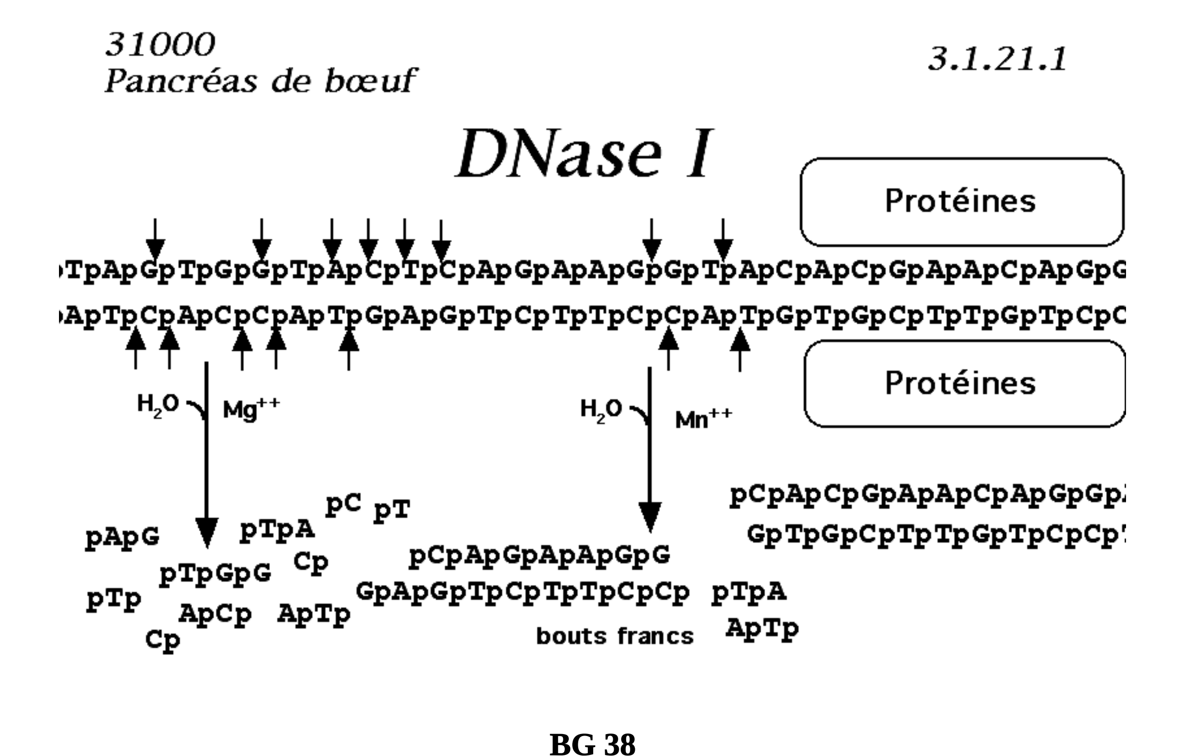


BG 37

- La ribonucléase H est une ribonucléase bactérienne qui intervient dans la maturation des RNA amorces qui servent à initier la réplication des plasmides.
- Elle est utilisée lors de la synthèse du deuxième brin d'un cDNA issu de la reverse transcription, afin de limiter les brins synthétisés en dehors de l'amorce spécifique (*self-primed strands*).



5.25 Désoxyribonucléase I

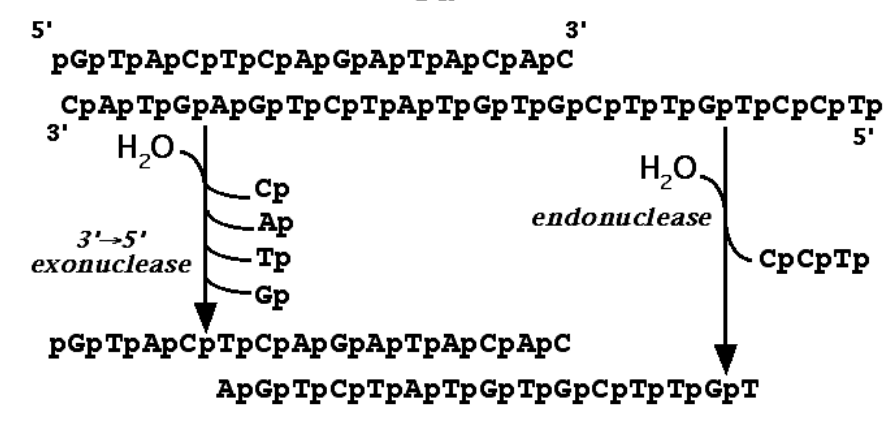


BG 38

- La désoxyribonucléase I est l'enzyme de la digestion des ADN chez les animaux. Elle hydrolyse les ADN double ou simple brin jusqu'à un mélange de nucléotides et d'oligonucléotides. Elle agit comme une endonucléase, préférentiellement sur les liaisons adjacentes aux nucléosides pyrimidiques. En présence de Mg^{++} elle hydrolyse les liaisons au hasard indépendamment de la séquence ; en présence de Mn^{++} elle devient plus dépendante de la séquence.
- La DNase pancréatique est utilisée pour introduire des brèches au hasard dans l'ADN double brin en vue d'un marquage par la DNA polymérase. Elle est également l'enzyme d'analyse des sites de liaisons protéine-DNA par la technique de *foot-printing*.



5.26 Nucléase BAL 31

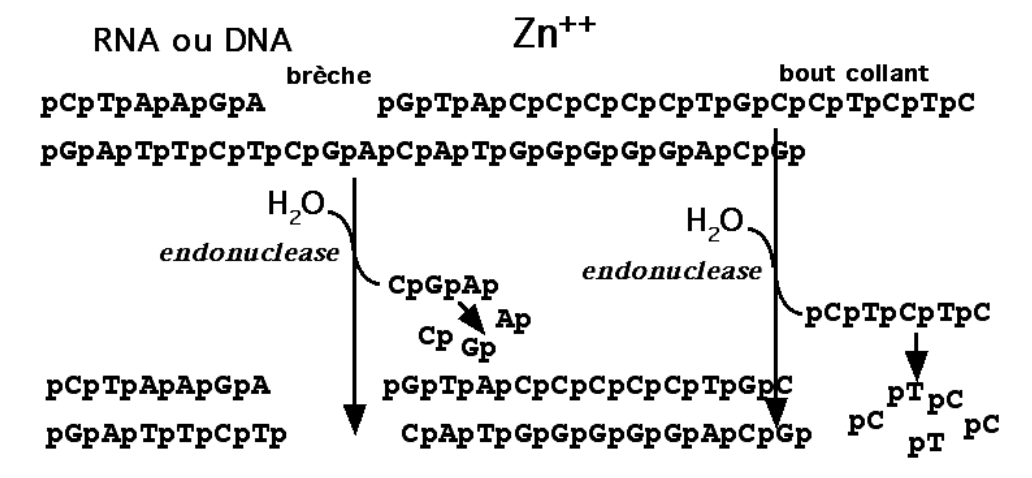


BG 39

- La nucléase BAL 31 est tout d'abord une exonucléase 3'→5' qui résorbe progressivement les extrémités 3' des ADN double brin. Sur les brins 5' sortants qu'elle isole elle agit ensuite comme une endonucléase pour détruire les fragments d'ADN simple brin. Elle est absolument dépendante de la présence de Ca^{++} comme cofacteur. Elle a peu d'activité sur les ARN.
- La nucléase BAL 31 est surtout utilisée pour le caractère progressif de son activité exonucléasique : on peut ainsi faire disparaître un à un les sites de restriction d'un ADN pour connaître l'ordre dans lequel ils sont présents dans la séquence de cet ADN.



5.27 Nucléase S1

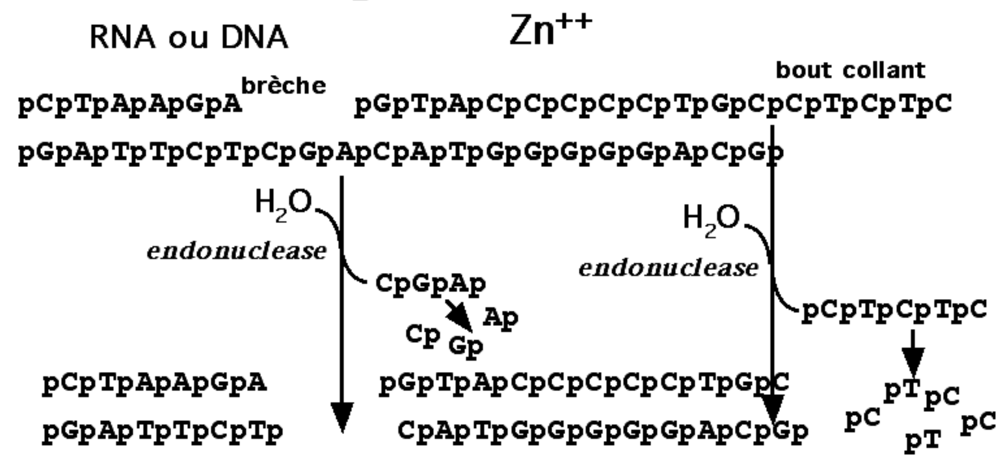


BG 40

- La nucléase S1 est une nucléase spécifique des ADN (ou ARN) simple brin, bien qu'à des concentrations élevées elle agisse aussi sur les hybrides. Elle agit en milieu acide, en présence d'ions Zinc.
- La nucléase S1 est utilisée :
  - — pour faire des bouts francs aux extrémités des fragments d'ADN double brin ;
  - — pour hydrolyser les fragments d'ADN simple brin au niveau des brèches (même s'il ne manque qu'une seule liaison) ;
  - — pour isoler les hybrides ADN:ARN lors de l'hybridation entre un gène et le cDNA correspondant ;
  - — pour ouvrir les épingles à cheveux formées lors de la synthèse des cDNA.



5.28 Nuclease de *mung-bean*



BG 41

- La nucléase des germes de haricot doré est aussi une nucléase spécifique des ADN (ou ARN) simple brin, bien qu'à des concentrations élevées elle agisse également sur les hybrides.
- Encore plus douce que la nucléase S1, elle respecte les brins d'ADN présentant des brèches d'une seule liaison (*nicks*).



5.29 Exonucléase de phage λ

Bacteriophageλ **3.1.11.3**

Bacteriophage λ exonucléase

dsDNA

pCpTpApApGpApGpCpTpGpTpApTpGpCpCpTpCpTpCpApC 3'

pGpApTpTpCpTpCpGpApCpApTpApCpGpGpApGpApGpTpGp 5'

5' exonuclease H_2O

Gp
Tp
Gp
Ap
Gp

pCpTpApApGpApGpCpTpGpTpApTpGpCpCpTpCpTpCpApC

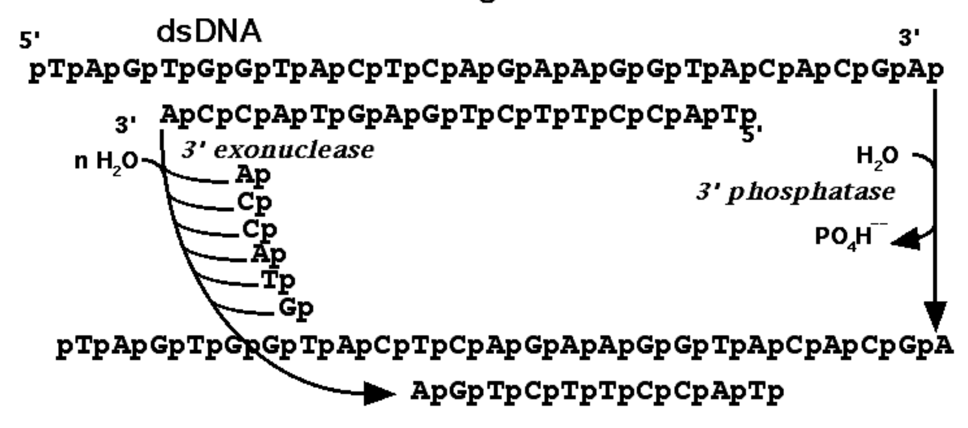
pGpApTpTpCpTpCpGpApCpApTpApCpGpGpAp

BG 42

- Les exonucléases sont des enzymes qui hydrolysent le premier ou le dernier nucléotide au bout d’un brin d’acide nucléique.
- L’exodésoxyribonucléase du bactériophage λ, qu’on rencontre aussi chez les bactériophages T4 ou T7, ainsi que chez les Mammifères (DNase IV), hydrolyse de préférence les extrémités 5’-phosphate des DNA double-brin en continuant vers le côté 3’. (5’$\rightarrow$3’ exonucléase). Elle produit des nucléosides 5’-phosphate.
- Elle est exprimée dans les cultures bactériennes infectées par ces phages.



5.30 Exonucléase III



BG 43

- Les exonucléases sont des enzymes qui hydrolysent le premier ou le dernier nucléotide au bout d'un brin d'acide nucléique.
- L'exodésoxyribonucléase III d'*Escherichia coli*, qu'on rencontre aussi chez *Haemophilus influenzae*, hydrolyse de préférence les extrémités 3'OH des DNA double-brin en remontant vers le côté 5'. (3'→5' exonucléase). Elle produit des nucléosides 5'-phosphate.
- Cette enzyme est aussi capable d'hydrolyser un radical phosphoryl lié à la fonction 3' alcool du ribose du dernier nucléotide (3'-phosphate terminal).




faculté de médecine
Pierre et Marie CURIE

Chapitre 6

Autres enzymes



6.1 β-galactosidase

540000
4 sous-unités

3.2.1.23

β-galactosidase

IPTG

X-Gal $\xrightarrow{H_2O}$ Galactose + 5-Br,4-Cl-Indol

BG 44

- La β-galactosidase est une enzyme d'*Escherichia coli*. Lorsque cette bactérie pousse sur un milieu riche en lactose, elle exprime la β-galactosidase qui est un des gènes de l'opéron lactose.
- La β-galactosidase hydrolyse spécifiquement les liaisons osides dans lesquelles l'anomère β du galactose engage son carbone réducteur.
- La β-galactosidase est utilisée comme gène « *reporter* » dans beaucoup de constructions de vecteurs. On introduit le promoteur de l'opéron lactose et le gène de la β-galactosidase dans un vecteur puis on transfecte des bactéries avec. Les bactéries utilisées sont dépourvues de l'opéron lactose (lac-).
- En faisant pousser les bactéries transfectées en présence d'un inducteur de l'opéron lactose (IPTG = isopropyl-thio-galactoside) on induit la synthèse de l'enzyme dans les colonies. En faisant pousser ces colonies en présence d'un substrat artificiel (X-Gal, composé incolore) on permet l'hydrolyse de ce substrat par l'enzyme qui libère le produit X (5Br,4Cl-indol, coloré en bleu).
- La présence de ce colorant donne la preuve de l'activité de la β-galactosidase, donc de son expression à partir du gène et par conséquent de la viabilité du vecteur transfecté dans les bactéries.



6.2 Protéinase K

Tritirachium album 3.4.21.64

Protéinase K

H_3N^+–His–Gly–Gly–Leu–Asn–Ala–Ala–Leu–Thr–COO^-

H_2O H_2O

H_3N^+–Asn–Ala–Ala–Leu–COO^-

H_3N^+–His–Gly–Gly–Leu–COO^- H_3N^+–Thr–COO^-

BG 45

- La protéinase K est une protéase végétale très active, même en présence de laurylsulfate (SDS) ou d'EDTA, et qui n'a aucun effet d'hydrolyse sur les acides nucléiques. Elle hydrolyse les protéines de toutes origines en quelques heures avec une préférence pour les liaisons peptidiques situées après les acides aminés hydrophobes (leucine, par exemple).
- La protéinase K est utilisée dans les techniques de préparation et de purification des acides nucléiques.




faculté de médecine
Pierre et Marie CURIE

Partie II

Préparation des acides nucléiques

Rappel des objectifs

- Définir[1] les termes suivants : oligonucléotide, sonde, ADN complémentaire.
- Décrire les gestes[2] qui permettent de faire l'extraction et la purification d'ADN génomique à partir d'un prélèvement de sang.
- Expliquer le principe de l'extraction de l'ARN messager d'un tissu et de la réalisation pratique d'une banque de cDNA.
- Décrire les gestes qui permettent de faire l'amplification par PCR d'un fragment d'ADN dont on possède les amorces.
- Expliquer le principe de la synthèse d'un oligonucléotide au moyen d'un synthétiseur et de la technique des phosphoramidites.
- Décrire les gestes qui permettent de faire la synthèse d'une queue de nucléotides à l'extrémité d'un fragment d'ADN.
- Expliquer le principe du dosage des acides nucléiques. Faire le calcul de la quantité d'acides nucléiques à partir des données brutes fournies par un spectrophotomètre, en tenant compte des dilutions et du volume du milieu considéré.
- Expliquer le principe de l'hybridation d'une sonde. Faire le calcul de la Tm de cette sonde à partir de sa séquence et d'une formule de calcul. Etablir les conditions des phases d'un programme de PCR utilisant deux amorces de Tm voisines. Expliquer les étapes des techniques d'hybridation moléculaire qui auront été pratiquées ou expliquées en cours à titre d'exemples : hybridation sur filtre, hybridation in situ, hybridation en solution.
- Décrire le principe technique et les gestes qui permettent de faire le marquage d'une sonde par un atome de ^{32}P ou par un nucléotide marqué.
- Expliquer le principe du séquençage d'un fragment d'ADN au moyen d'un séquenceur et de la technique des *dye primers* ou des *dye terminators* ; décrire et interprèter les résultats[3] apparaissant sur l'écran de l'ordinateur après un séquençage automatique.

---

1. **Définir** : préciser dans une phrase concise l'essence d'un objet ou les limites d'un concept en excluant toute notion étrangère et en comprenant toutes les variations possibles de l'objet ou du concept cerné.
2. **Expliquer le principe**, **décrire ou mimer les gestes** ou **faire un schéma explicatif** : préciser le motif de chacun de ces gestes. Etre capable de déceler une erreur qualitative dans la description d'un protocole.

Au niveau des travaux pratiques, les connaissances techniques théoriques étant acquises, on devra être capable de passer à la manipulation sans avoir à apprendre que la partie proprement manuelle.

3. **Interpreter les résultats** : à partir des données affichées par la machine, après la validation technique de la séquence, savoir décrire la structure de l'acide nucléique séquençé et reconnaître sur cette séquence des lésions moléculaires dont l'interprétation ne nécessite la levée d'aucune ambiguïté.

- Expliquer les principe de la recherche des éléments cis-régulateurs sur la séquence d'un promoteur (*foot printing*).
- Expliquer le principe des réactions permettant le dosage des ARN transcrits : PCR quantitative, protection à la Rnase.

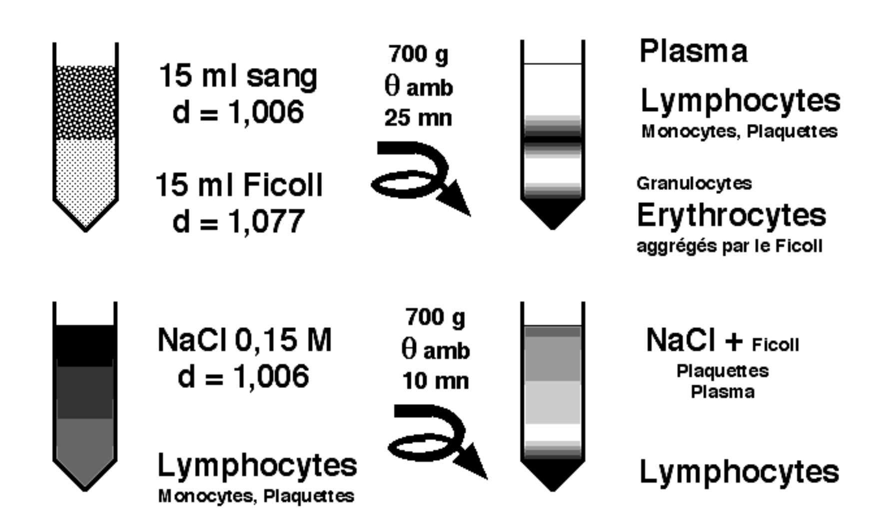


Chapitre 7

Extraction et purification



7.1 Purification des lymphocytes

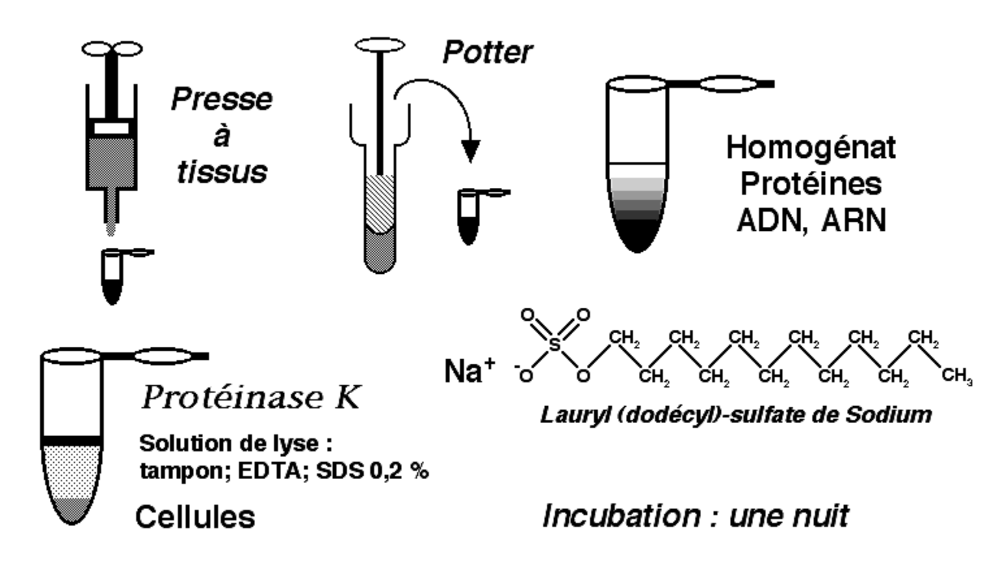


BG 46

- Le prélèvement de sang veineux au pli du coude est la technique la plus employée pour obtenir de l’ADN génomique. Le sang est recueilli sur EDTA, chélateur des ions divalents, cofacteurs des nucléases.
- Pour faciliter l’extraction de l’ADN on isole du sang total un culot de cellules nucléées, les lymphocytes.
- Le sang est d’abord déposé en gradient de densité dans un tube à centrifuger, au-dessus d’une couche de polyfluorocarbone liquide (Ficoll ®) de densité 1,077. Le plasma flotte au dessus de ce produit car sa densité est de 1,006. Puis on centrifuge ce gradient pendant 25 min à 700 g à la température ambiante.
- Les lymphocytes et d’autres cellules blanches sont recueillies à la limite supérieure de la couche de Ficoll, tandis que les autres cellules plus denses ont sédimenté au fond du tube.
- Les lymphocytes sont enfin lavés dans du NaCl 0,15 M et recentrifugés 10 min dans les mêmes conditions. Ils forment après cette deuxième centrifugation un culot blanc au fond du tube.



7.2 Homogénéisation des tissus, déprotéinisation



BG 47

- L'extraction de l'ADN à partir des tissus se fait par broyage dans une presse à tissus, suivie d'une homogénéisation dans un tube de Potter.
- L'homogénat qu'on obtient est un mélange de débris de membranes (microsomes), de nombreuses particules subcellulaires (noyaux, mitochondries) et de molécules en solution (cytoplasme).
- On dissous cet homogénat avec une solution de lyse (tampon sans pression osmotique, faisant éclater les membranes fermées) en présence d'EDTA et de laurylsulfate de Sodium (SDS) qui inhibent les nucléases.
- On incube enfin le milieu avec la protéinase K, pour détruire les protéines en présence du détergent (SDS), à la température ambiante et pendant toute la nuit.



7.3 Extraction de l’ADN

Extraction de l'ADN

Cellules lysées ou tissu homogénéisé protéines, ADN, ARN

Phénol distillé saturé Tris d = 1,07

Agiter

10000 g θ 4°C 10 mn

Surnageant ADN, ARN

Protéines

Phénol

Surnageant ADN, ARN

$CHCl_3$ / Isoamyl-OH v/v 24:1 d = 1,45

10000 g θ 4°C 10 mn

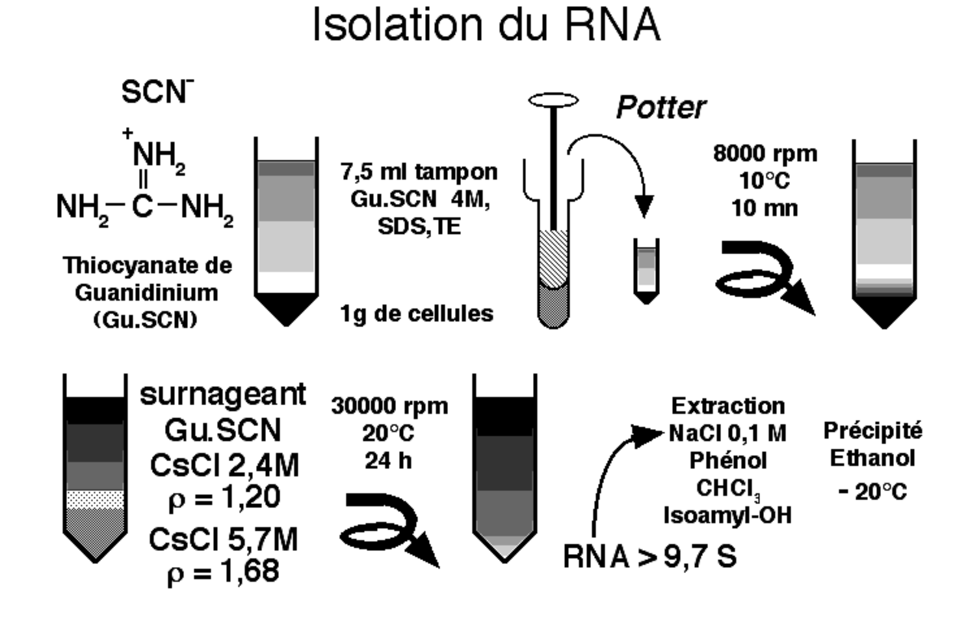
Surnageant ADN, ARN en solution

BG 48

- Les solutions d’acides nucléiques et de protéines issues des préparations précédentes sont purifiées par addition de phénol saturé de tampon Tris-EDTA.
- Après agitation douce du mélange, on sépare les phases par une centrifugation de 10 min sous 10000 g à la température de 4°C. A la sortie, on distingue une phase aqueuse surnageante contenant les acides nucléiques en solution, une phase organique au fond du tube : phénol + lipides et à l’interface un « gateau » de protéines précipitées.
- On recueille la phase aqueuse avant de la laver par un mélange de chloroforme ($CHCl_3$) et d’alcool isoamylique dans un rapport 24:1 (volume:volume). On recueille à nouveau une phase aqueuse contenant l’ADN en solution et une phase organique (24:1) contenant des restes de phénol, de protéines et de lipides.



7.4 Isolation de l’ARN



BG 49

- Pour extraire l’ARN d’un tissu ou d’une suspension de cellules, il faut impérativement inhiber toute trace de RNase.
- L’homogénéisation dans le tube de Potter se fait dans un tampon Tris-EDTA en présence de thiocyanate de guanidinium 4 M et de SDS. Les débris cellulaires sont sédimentés en 10 min à 10 °C.
- Le surnageant contenant les acides nucléiques est déposé sur un gradient de densité comprenant au fond du tube une solution très dense de CsCl 5,7 M surmontée d’une couche intermédiaire de CsCl 2,4 M. Ce gradient est ultracentrifugé durant 24 heures à 30000 tours/min à la température ambiante.
- On peut alors recueillir un culot d’ARN de masses moléculaires élevées qui seront soumis à l’extraction par le phénol et le chloroforme.



7.5 Techniques de purification (tableau)

Purification

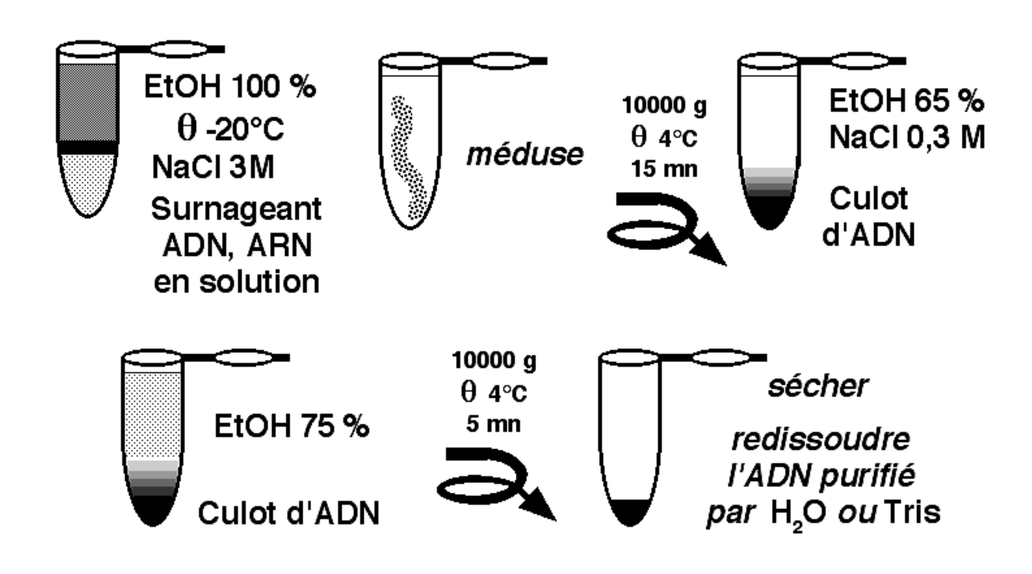
- **Gel dénaturant**
  - acrylamide 15 % urée 10 M
- **Chromatographie sur gel de silice**
  - adsorption : solvant polaire (tampon)
  - élution : solvant apolaire (eau + alcool)
- **Précipitation par l'éthanol froid**
  - pour des longueurs supérieures à 15 pb

BG 50

- Il existe de nombreuses techniques de purification des acides nucléiques.
- On peut purifier les ADN par chromatographie dans une colonne de gel en présence d'un dénaturant ou bien par chromatographie d'adsorption.
- Le plus souvent on se contente de multiples précipitations par l'éthanol qui suffisent à débarrasser l'ADN des protéines contaminantes et des enzymes.



7.6 Purification par l'éthanol

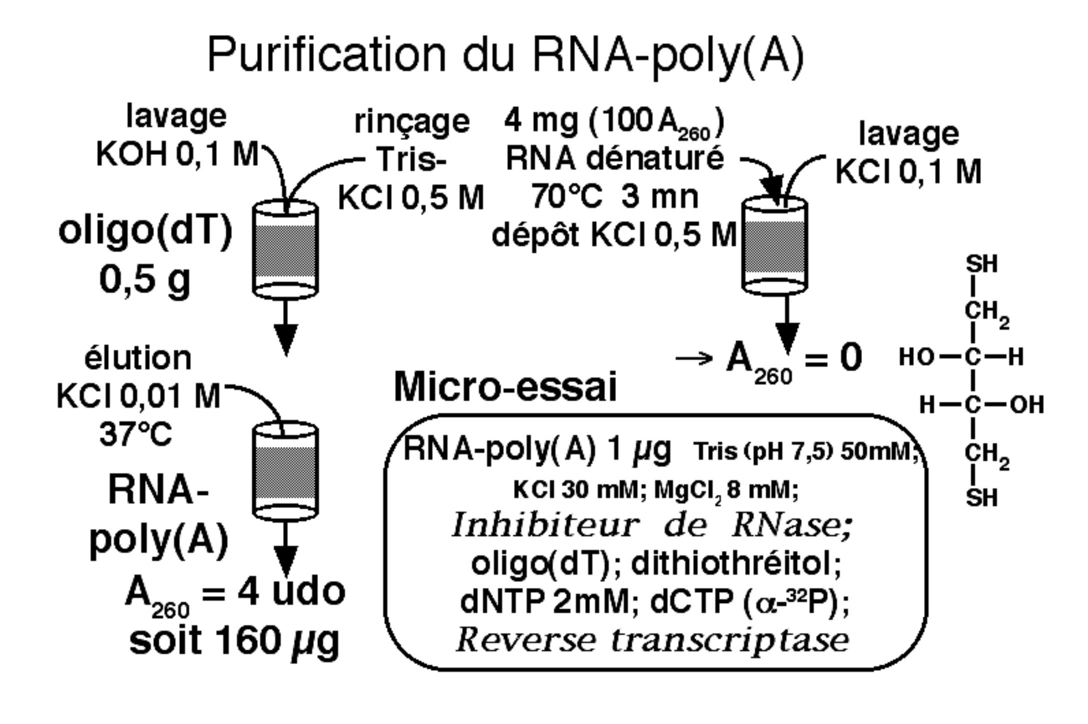


BG 51

- La purification d'une solution d'ADN se fait le plus souvent par précipitations répétées dans l'alcool.
- Une solution d'ADN, portée à haute force ionique par addition de NaCl 3 M (1/10 du volume), est additionnée d'un large excès d'éthanol absolu à -20°C. Au bout de quelques minutes au maximum, on voit apparaître un précipité blanchâtre, translucide et filamenteux (la « méduse ») constitué de longs filaments d'ADN précipité.
- On recueille ce précipité par centrifugation à 10000 g pendant un quart d'heure environ. Pour laver l'ADN on resuspend le précipité dans de l'éthanol à 75 % toujours à -20°C. puis on centrifuge à nouveau. Le culot de cette dernière centrifugation est égoutté, puis séché sous vide.
- On peut conserver l'ADN à sec ou le redissoudre immédiatement soit dans de l'eau pure stérile, soit dans un tampon Tris-EDTA.



7.7 Purification du RNA-poly(A)



BG 52

- Parmi les acides ribonucléiques extraits des cellules, la classe la plus étudiée est celle des ARN messagers. Ils se caractérisent par la présence du côté 3’-terminal d’une longue queue poly(A) synthétisée à la phase post-transcriptionnelle.
- On utilisera une chromatographie d’affinité entre cette queue poly(A) et une colonne dont la phase fixe est pourvue de fragments d’oligo(dT) de quelques dizaines de nucléotides qui peuvent s’hybrider avec les RNA poly(A).
- Après avoir lavé la colonne en milieu alcalin (potasse), on charge les RNA totaux à haute force ionique (KCl 0,5 M) puis on rince la colonne avec du KCl 0,1 M en surveillant l’absorbance de l’éluat à 260 nm pour s’assurer de l’élimination des RNA non retenus par la colonne (rRNA, tRNA,...). Enfin, on élue la fraction retenue par la colonne en abaissant la concentration de KCl à 0,01 M et en élevant la température.
- Un micro essai avec la reverse transcriptase permet de s’assurer de la qualité de ce RNA poly(A) comme matrice pour la synthèse de cDNA.

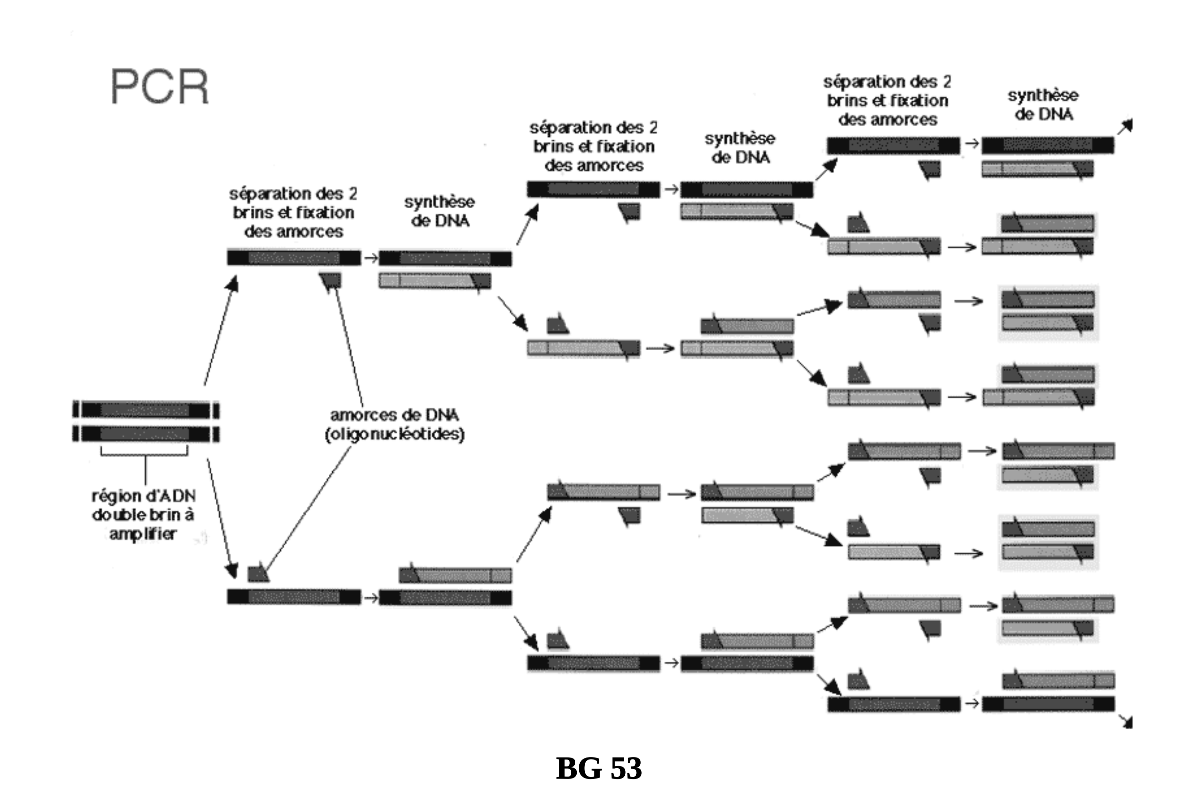


Chapitre 8

Synthèse des polynucléotides



8.1 Polymerase Chain Reaction (PCR)

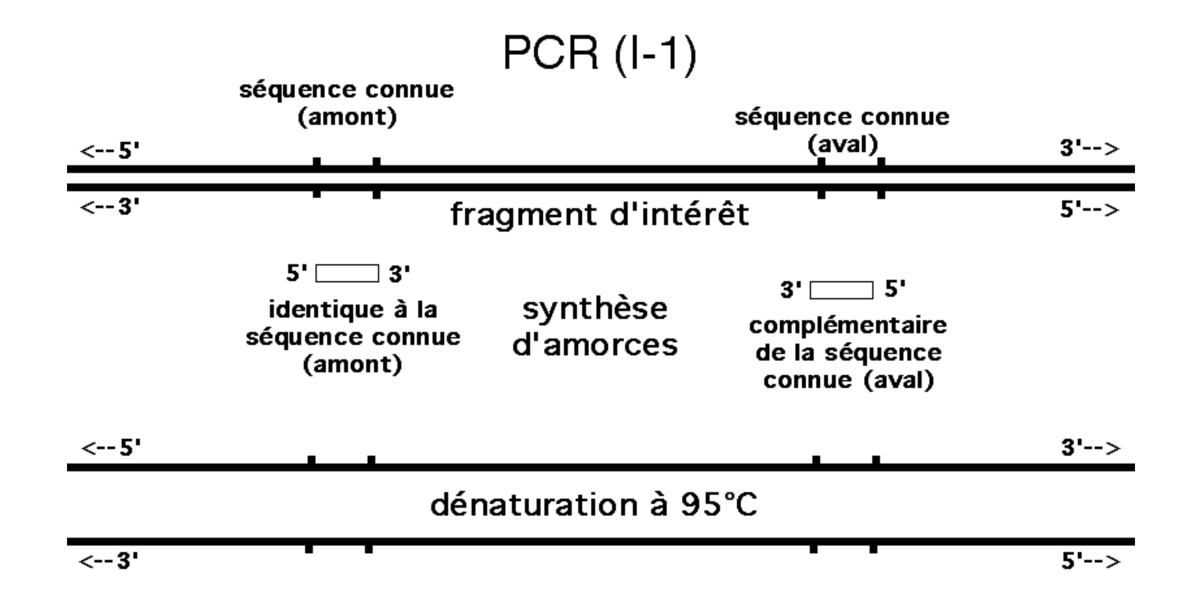


BG 53

- La PCR (*polymerase chain reaction*) permet d'amplifier spécifiquement une région de DNA double brin de quelques centaines de paires de bases. Ce DNA doit d'abord être séparé en simples brins (dénaturation à 95°C).
- On ajoute au DNA de départ une large quantité d'amorces (oligonucléotides synthétiques complémentaires des deux extrémités de la région à amplifier) qui vont s'hybrider à 50-60°C environ avec la séquence complémentaire sur chacun des brins de DNA, et les quatre dNTP qui serviront de substrats.
- On soumet le tout à l'activité d'une DNA polymérase (Taq polymerase) qui synthétise à 72°C un brin complémentaire à partir du 3'OH de l'amorce hybridée. On obtient quatre brins de DNA.
- On recommence à dénaturer ces 4 brins, puis on les laisse s'hybrider avec les amorces (toujours en excès) et la polymérase entre en action pour aboutir à 8 brins de DNA.
- On recommence à dénaturer ces 8 brins, puis on les laisse s'hybrider avec les amorces (toujours en excès) et la polymérase entre en action pour aboutir à 16 brins de DNA.
- Et ainsi de suite 30 à 40 fois, ce qui aboutit à 34359738368 brins de DNA (34 milliards), ce qui représente une quantité suffisante pour étudier le fragment de DNA amplifié.



8.2 PCR (I-1)

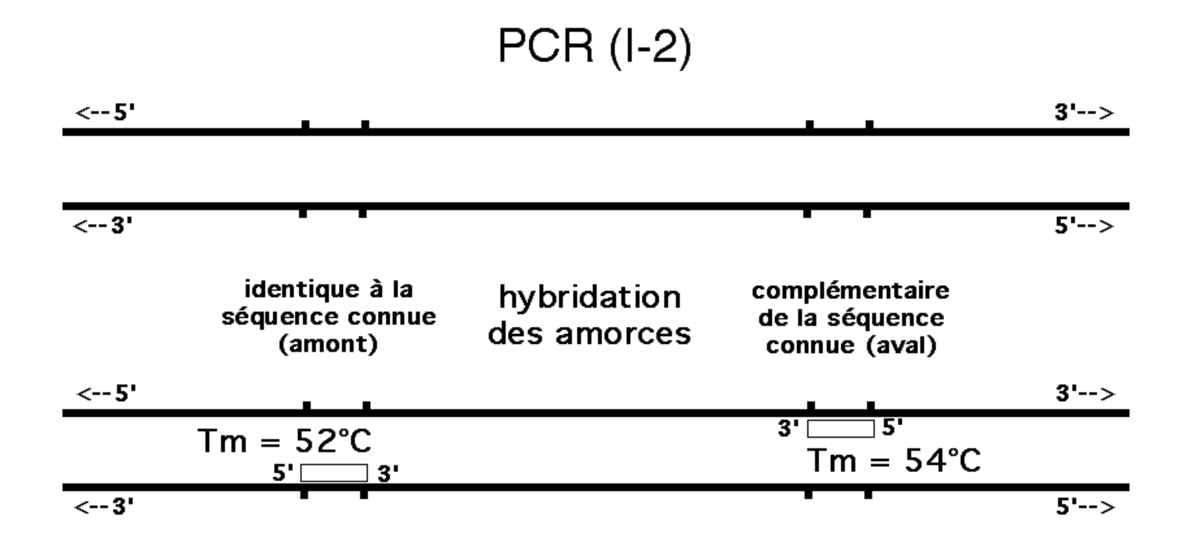


BG 53/1

- La PCR (*polymerase chain reaction*) est destinée à synthétiser un grand nombre de fragments d'ADN identiques à un fragment d'intérêt afin de pouvoir le caractériser. Ce fragment à multiplier est situé dans un ADN génomique ou dans un cDNA. Il est indispensable que soient connues les séquences aussi exactes que possible de deux petits oligonucléotides (20 à 30 nt) situés aux deux extrémités du fragment d'intérêt (côté 5' ou amont ; côté 3' ou aval).
- On synthétise ces oligonucléotides pour servir d'amorces à la réaction. L'amorce côté 5' est identique à la séquence connue en amont. L'amorce côté 3' est antiparallèle et complémentaire de la séquence connue en aval, c'est à dire identique à celle de l'autre brin à ce niveau.
- Pour commencer la réaction on dénature l'ADN en le chauffant à 95° plusieurs minutes, pour séparer les deux brins.



8.3 PCR (I-2)

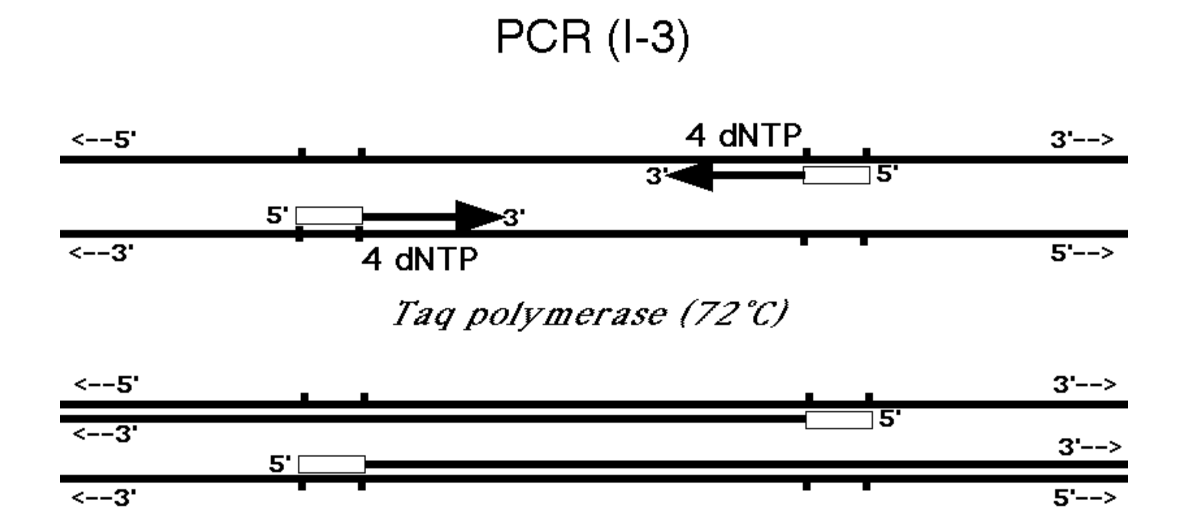


BG 53/2

- Chacune des amorces amont et aval est complémentaire d'une séquence connue et on peut déterminer la Tm de leur hybridation : ici, 52°C pour la séquence amont et 54°C pour la séquence aval.
- En refroidissant la température vers 50°C (un peu en-dessous des Tm des deux amorces) les fragments d'ADN dénaturé vont pouvoir s'hybrider avec des séquences complémentaires. La concentration des amorces étant beaucoup plus élevée que celle des brins d'ADN, chaque brin d'ADN s'hybridera avec une amorce qui lui est complémentaire.
- Ainsi le brin du bas, orienté 5'→3' de droite à gauche, s'hybridera avec l'amorce amont, orientée 5'→3' de gauche à droite et le brin du haut, orienté 5'→3' de gauche à droite, s'hybridera avec l'amorce aval, orientée 5'→3' de droite à gauche.



8.4 PCR (I-3)

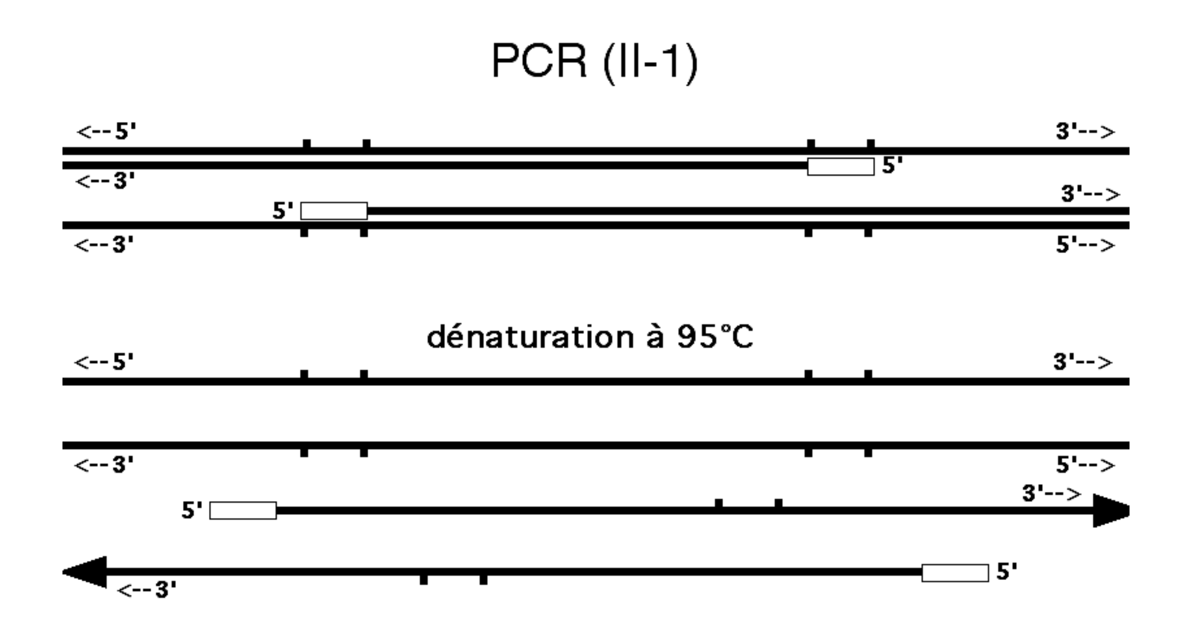


BG 53/3

- A 95°C durant la phase de dénaturation, puis à 50° C durant la phase d'hybridation, la Taq polymérase était inactive parce que trop éloignée de sa température optimale d'activité (75 à 80°C).
- En remontant la température à 72°C, l'enzyme commence son activité. Comme toutes les DNA polymérases elle nécessite un ADN double brin avec une extrémité 3'OH rentrante pour initier la polycondensation des désoxyribonucléotides, elle lit le brin modèle de 3' vers 5' et elle construit le nouveau brin de 5' vers 3' qui se prolonge autant qu'il existe un brin complémentaire pour servir de modèle.
- Le nouveau brin d'ADN est formé à partir de l'amorce amont (côté 5') et se prolonge vers la droite (côté 3'). L'autre brin d'ADN est formé à partir de l'amorce aval (côté 5') et se prolonge vers la gauche (côté 3'). La synthèse n'est limitée que par le temps d'incubation (environ 1000 nucléotides par minute).



8.5 PCR (II-1)

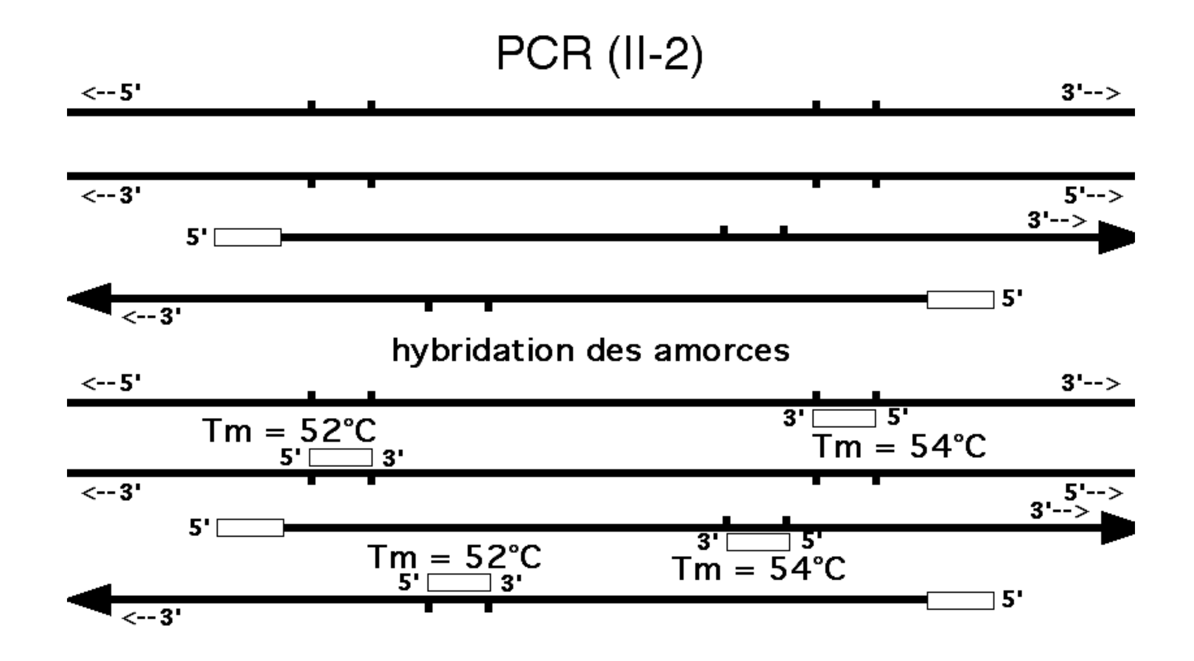


BG 53/4

- Au cours de la phase d'élongation chacun des brins du fragment d'intérêt a été répliqué une fois, la nombre de copies a donc doublé.
- Le deuxième cycle commence par la phase de dénaturation pendant laquelle on dénature à nouveau l'ADN en le chauffant à 95° plusieurs minutes, pour séparer les deux brins d'origine des brins nouvellement synthétisés.



8.6 PCR (II-2)

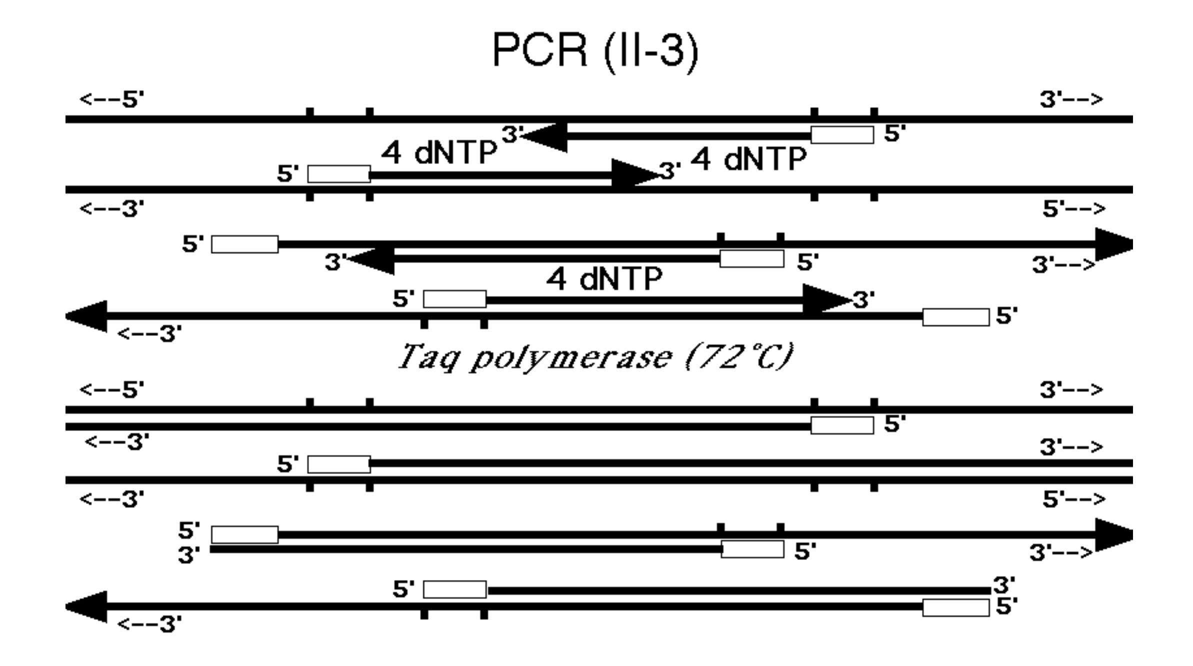


BG 53/5

- Les brins d'ADN nouvellement synthétisés ont répliqué le fragment d'intérêt et aussi la séquence complémentaire de l'amorce qui le suit.
- En refroidissant la température vers 50°C (un peu en-dessous des Tm des deux amorces) les fragments d'ADN dénaturé et les fragments d'ADN nouvellement synthétisés vont pouvoir s'hybrider à nouveau avec les amorces.
- Ainsi les brins orientés 5'→3' de droite à gauche s'hybrideront avec l'amorce amont, orientée 5'→3' de gauche à droite et les brins orientés 5'→3' de gauche à droite s'hybrideront avec l'amorce aval, orientée 5'→3' de droite à gauche.



8.7 PCR (II-3)

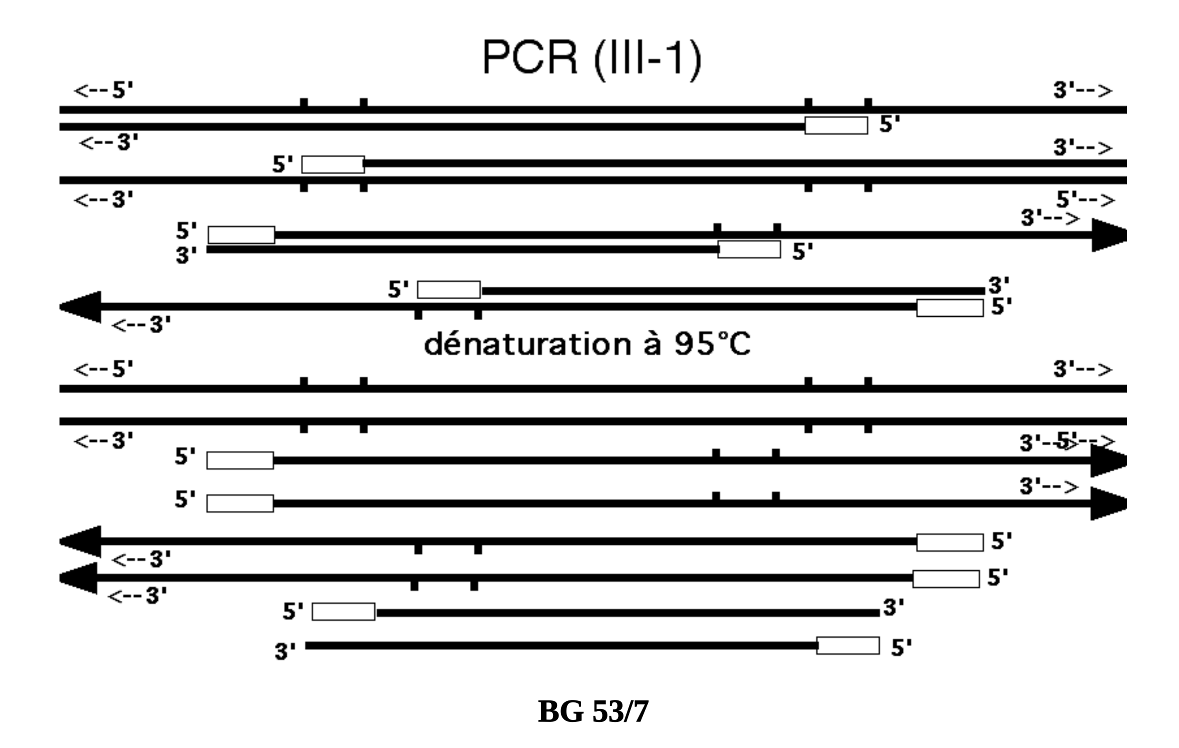


BG 53/6

- La Taq polymérase, thermorésistante, a été inactivée mais pas dénaturée durant la phase de dénaturation de l'ADN.
- En remontant la température à 72°C, l'enzyme recommence donc son activité. Elle catalyse la polycondensation des désoxyribonucléotides, qui se prolonge autant qu'il existe un brin complémentaire pour servir de modèle.
- Le nouveau brin d'ADN formé à partir de l'amorce amont (côté 5') et se prolonge vers la droite (côté 3'). L'autre brin d'ADN formé à partir de l'amorce aval (côté 5') et se prolonge vers la droite (côté 3'). Sur les brins d'ADN (en jaune) synthétisés à partir des amorces au cycle précédent, la synthèse sera limitée au côté 5' de l'amorce opposée, faute de modèle pour poursuivre. Seul le fragment d'intérêt sera répliqué sur ces nouveaux fragments.



8.8 PCR (III-1)

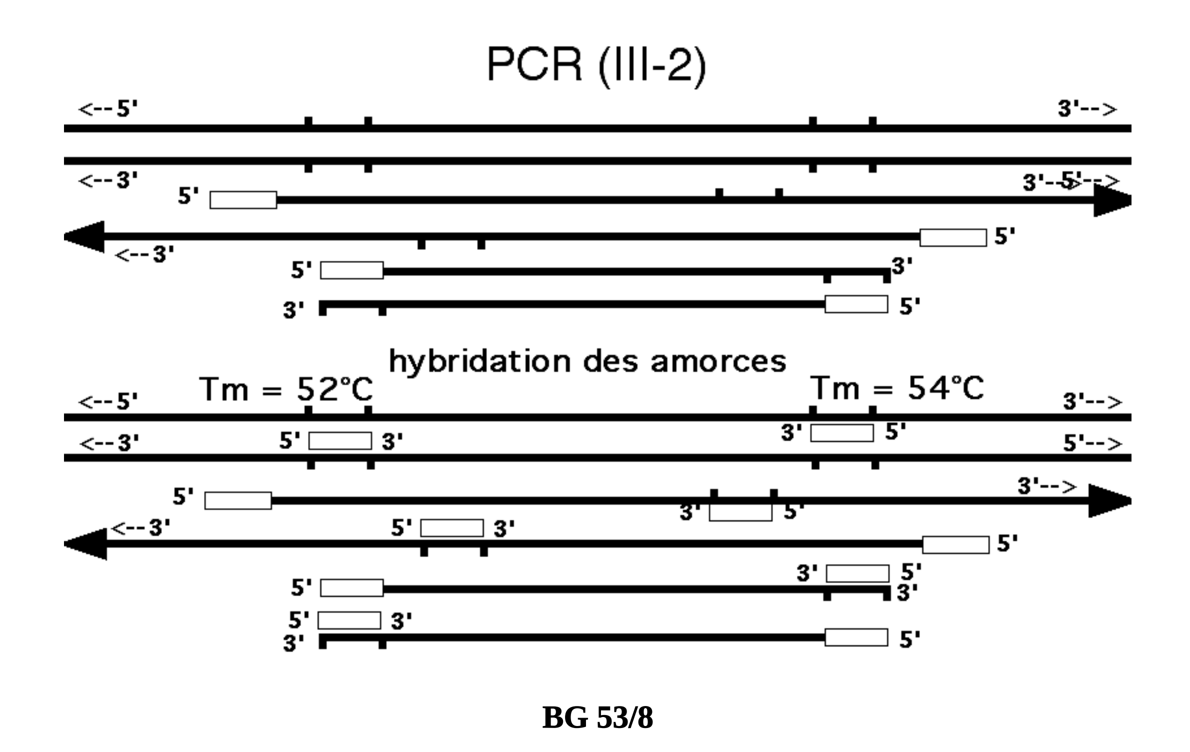


BG 53/7

- Au cours de la précédente phase d'élongation chacun des brins d'ADN contenant la séquence du fragment d'intérêt a été répliqué une fois, la nombre de copies a donc encore doublé. A chaque cycle ce nombre doublera à nouveau et l'amplification se poursuit comme la suite des puissances de 2 : 2^2, 2^3, 2^4, 2^5, 2^6, 2^7,... à la fin du 35ème cycle, il y aura 2^{35} copies soit plus de 34 milliards de copies !
- Le cycle suivant commence par la phase de dénaturation pendant laquelle on dénature à nouveau l'ADN en le chauffant à 95° plusieurs minutes. A ce stade, il reste encore les deux ADN de départ, quelques ADN limités par une seule des séquences amorces et les séquences, de plus en plus nombreuses, limitées par les séquences des deux amorces.



8.9 PCR (III-2)

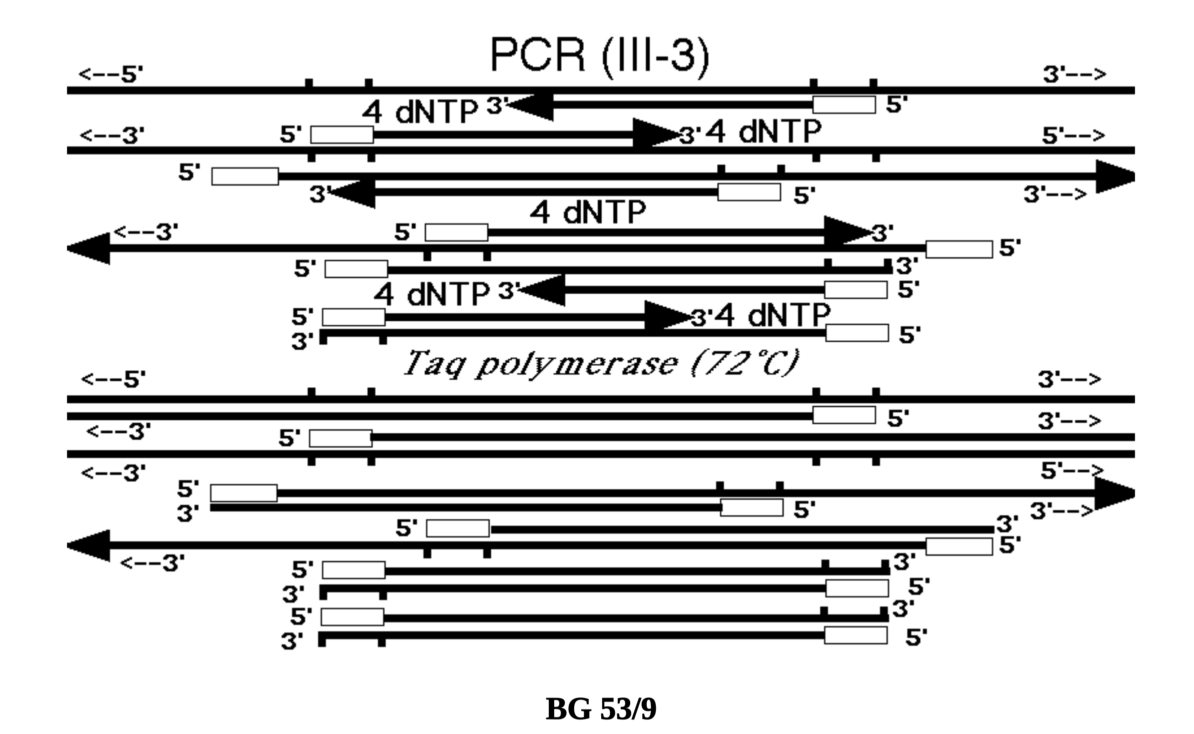


BG 53/8

- En refroidissant la température vers 50°C (un peu en-dessous des Tm des deux amorces) les fragments d'ADN dénaturé et les fragments d'ADN nouvellement synthétisés vont pouvoir s'hybrider à nouveau avec les amorces.
- Ainsi les brins orientés 5'→3' de droite à gauche s'hybrideront avec l'amorce amont, orientée 5'→3' de gauche à droite et les brins orientés 5'→3' de gauche à droite s'hybrideront avec l'amorce aval, orientée 5'→3' de droite à gauche.



8.10 PCR (III-3)

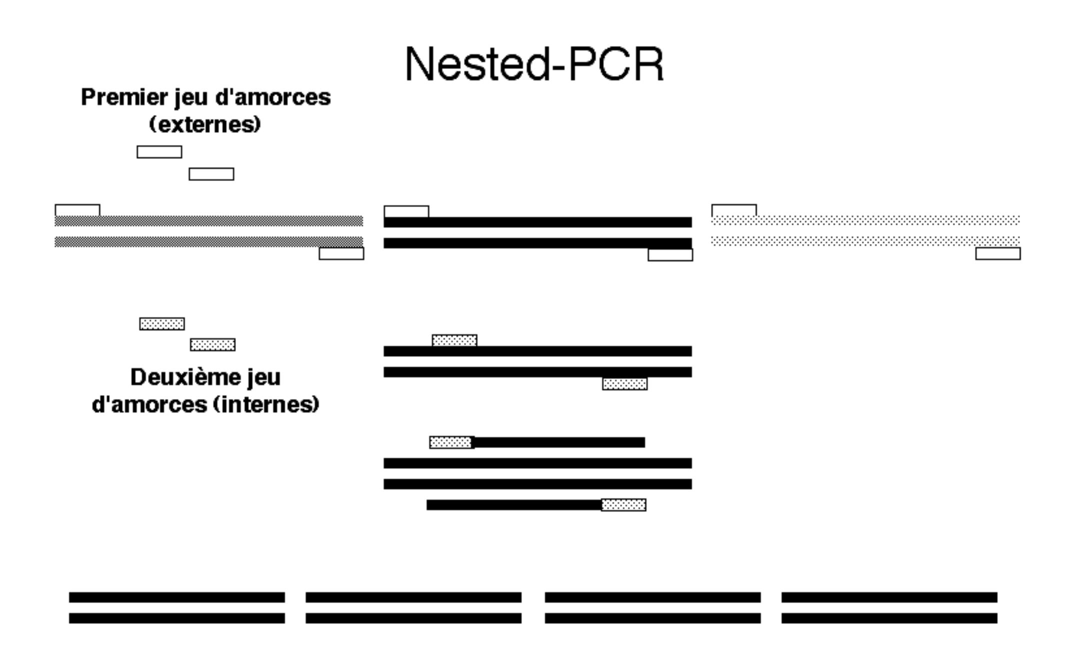


BG 53/9

- En remontant la température à 72°C, l'enzyme recommence encore son activité. Elle catalyse la polycondensation des désoxyribonucléotides, qui se prolonge autant qu'il existe un brin complémentaire pour servir de modèle.
- Le nouveau brin d'ADN formé à partir de l'amorce amont (côté 5') et se prolonge vers la droite (côté 3'). L'autre brin d'ADN formé à partir de l'amorce aval (côté 5') et se prolonge vers la gauche (côté 3'). Sur les brins d'ADN (en jaune) synthétisés à partir des amorces au cycle précédent, la synthèse sera limitée au côté 5' de l'amorce opposée, faute de modèle pour poursuivre. Seul le fragment d'intérêt sera répliqué sur ces nouveaux fragments.
- Le nombre de copies de l'ADN d'origine avec une seule amorce-limite va continuer à croître arithmétiquement et le nombre de copies avec deux amorces-limites continuera à croître exponentiellement, jusqu'à ce que ces dernières, soient pratiquement pures dans l'ADN amplifié final.



8.11 Nested-PCR



BG 54

- Les amorces spécifiques sont essentielles dans les techniques de PCR car elles seules vont déterminer le fragment à amplifier. Lorsqu'il existe des séquences homologues cette spécificité n'existe plus et la technique doit être modifiée pour mieux cibler le fragment à amplifier.
- On utilise deux couples d'amorces. Des amorces externes dont la spécificité peut permettre d'amplifier à la fois un petit nombre de séquences homologues et des amorces internes permettant d'amplifier spécifiquement une des séquences révélées par le couple d'amorces externes.
- Cette technique dite de *nested*-PCR (PCR nichée) se pratique comme la PCR avec le couple d'amorces externes pour dix cycles d'amplification. A cette étape on ajoute un excès d'une ou de deux amorces internes avant de prolonger l'amplification pour encore 25 cycles environ. Cette procédure permet le plus souvent d'obtenir l'amplification spécifique d'un seul fragment.



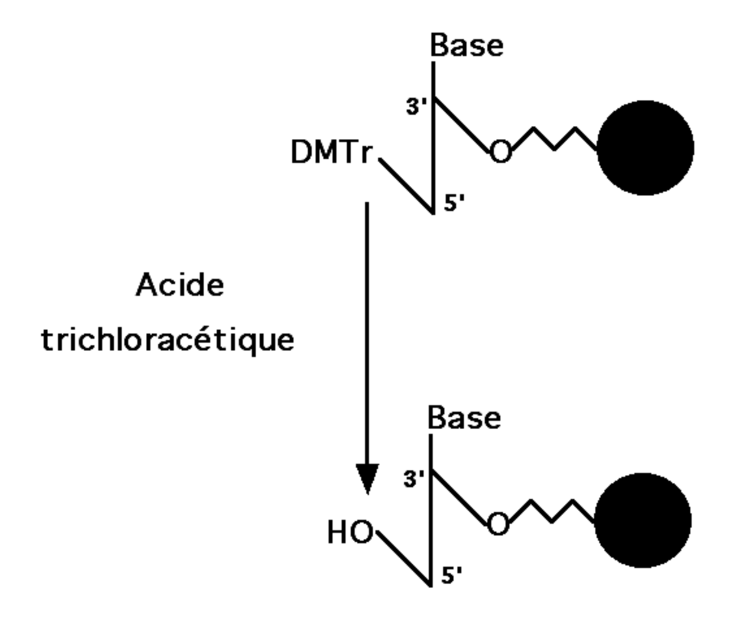
8.12 Phosphoramidites : dA-3'-support protégée

BG 55

- La synthèse des oligonucléotides se fait sur la phase fixe d'une colonne parcourue par un flux de liquide contenant successivement les réactifs nécessaires.
- Contrairement à ce qui se passe dans la nature, la synthèse chimique des oligonucléotides se fait de 3' vers 5'.
- Sur les particules solides de la colonne est fixé au départ un nucléotide qui sera le dernier de la séquence à synthétiser (extrémité 3'OH terminale). Ce nucléotide est lié par sa fonction alcool secondaire en 3' à une fonction amine du support, par l'intermédiaire d'un acide succinique pour permettre la libre rotation de l'ensemble par rapport au support.
- Pour permettre de diriger la synthèse le nucléotide terminal a toutes ses fonctions bloquées par des radicaux empêchant les réactions prématurées ou parasites :

  — radical diméthoxytrityl (DMTr), pour l'alcool primaire en 5' du désoxyribose
  — acide benzoïque, pour les fonctions amines de l'adénine ou de la cytosine
  — acide isobutyrique pour la fonction amine de la guanine.

8.13 Phosphoramidites : détritylation



BG 55/1

- Avant de faire la liaison du deuxième nucléotide (qui sera l'avant-dernier de l'oligonucléotide final) on enlève le radical DMTr qui protège l'extrémité 5'OH du nucléotide de départ.
- Cette détritylation est obtenue en faisant passer sur la colonne une solution d'acide trichloracétique (TCA) qui hydrolyse la liaison éther entre le DMTr et la fonction 5'OH du désoxyribose.



8.14 Phosphoramidites : addition d’un nucléotide

BG 55/2

- L’addition d’un nouveau nucléotide est faite en chargeant la colonne d’une solution contenant le nouveau nucléotide avec des fonctions protégées :
  - — par le DMTr pour la fonction 5’OH
  - — par les réactifs (benzoate, isobutyrate) protégeant les fonctions des bases azotées
  - — le phosphore lié à la fonction alcool en 3’ est un phosphoramidite dont la fonction amine est substituée par deux radicaux isopropyl et la fonction acide par un radical méthyl.
- En présence de tétrazole, le nouveau nucléotide chargé voit sa liaison phosphoamide hydrolysée et la fonction acide ainsi apparue va estérifier la fonction alcool déprotégée du dernier nucléotide précédemment fixé à la colonne.
- La liaison phosphodiester 3’$\rightarrow$5’ est créée mais le phosphore est encore réduit (trivalent).



8.15 Phosphoramidites : blocage des supports

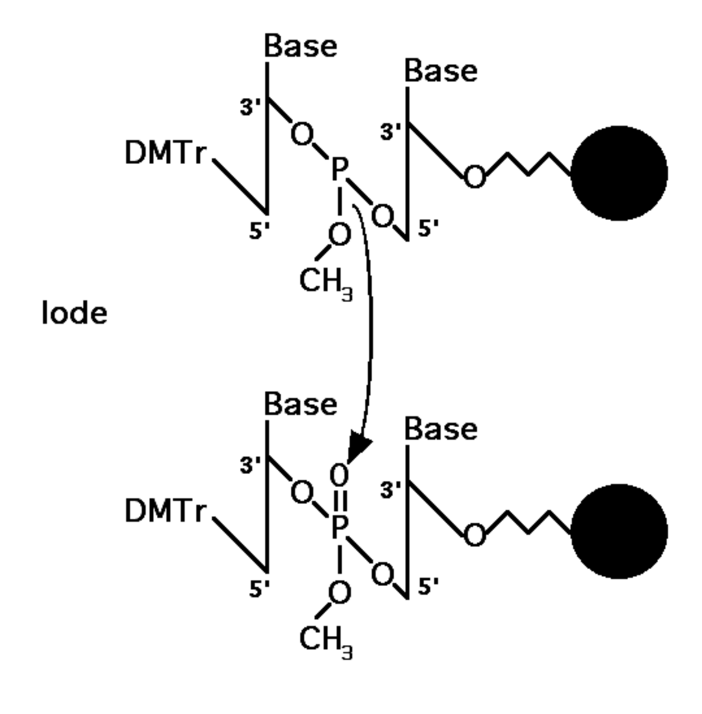
Base
3'
HO
5'
Anhydride
acétique
Base
3'
$H_3C-C(=O)-O$
5'

BG 55/3

- Toutes les fonctions alcool primaires de la colonne peuvent réagir avec les fonctions acides libérées par le tétrazole. En principe, seules les fonctions 5'OH terminales du dernier nucléotide lié à la colonne ont été déprotégées et peuvent réagir de la sorte. Le rendement de cette réaction est presque total (> 99 %).
- Afin qu'aucune erreur de séquence ne soit possible par la suite, on passe de l'anhydride acétique sur la colonne pour bloquer à nouveau les fonctions alcool estérifiables qui par accident n'auraient pas été liées au phosphoramidite à l'étape précédente. Les oligonucléotides qui n'auraient pas réagi avec le nouveau nucléotide sont ainsi écartés définitivement de la suite du cycle de synthèse.



8.16 Phosphoramidites : oxydation



BG 55/4

- Le phosphore trivalent de la liaison phosphodiester est ensuite oxydé par l'iode en phosphore pentavalent (phosphate).
- La seule fonction acide de ce phosphate non liée aux désoxyriboses reste bloquée par le radical méthyl.
- La colonne porte maintenant un nucléotide de plus dont les fonctions sont bloquées.
- L'introduction du nucléotide suivant sera précédée de la déprotection de la fonction alcool primaire en 5', présentement bloquée par un DMTr, et le cycle recommence : addition d'un nucléotide, blocage des supports, oxydation, déprotection, etc, autant de fois qu'il faudra de nucléotides pour compléter la séquence demandée.



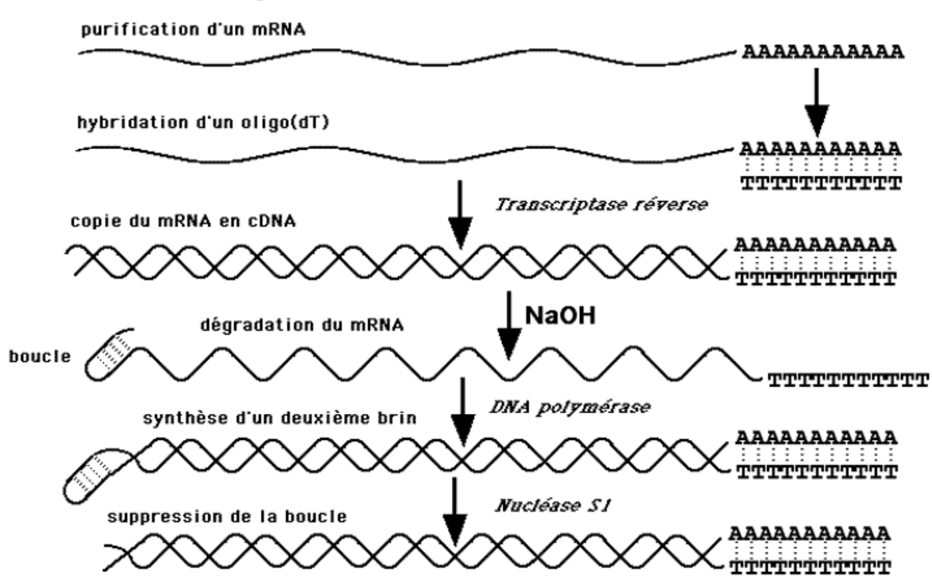
8.17 Phosphoramidites : déprotection

BG 55/6

- Lorsque le dernier cycle de synthèse a eu lieu, l'oligonucléotide est encore lié à la colonne par son extrémité 3' et toutes ses fonctions réactives sont protégées.
- Afin de donner au produit la structure d'un ADN normal, on déprotegera progressivement :
  - — acide trichloracétique pour déprotéger la fonction 5'OH initiale,
  - — thiophénol pour transformer les phosphotriesters en phosphodiesters en enlevant les groupes méthyl,
  - — ammoniaque concentrée à 25 °C pour détacher l'oligonucléotide de son support et enlever une partie des radicaux acétyl de protection,
  - — le reste des radicaux acétyl et les radicaux benzoyl sont détachés par l'ammoniaque concentrée à 55 °C.
- Le produit final est habituellement purifié sur colonne ou dans un gel pour éliminer les produits de synthèses incomplètes.



8.18 Synthèse d'un cDNA

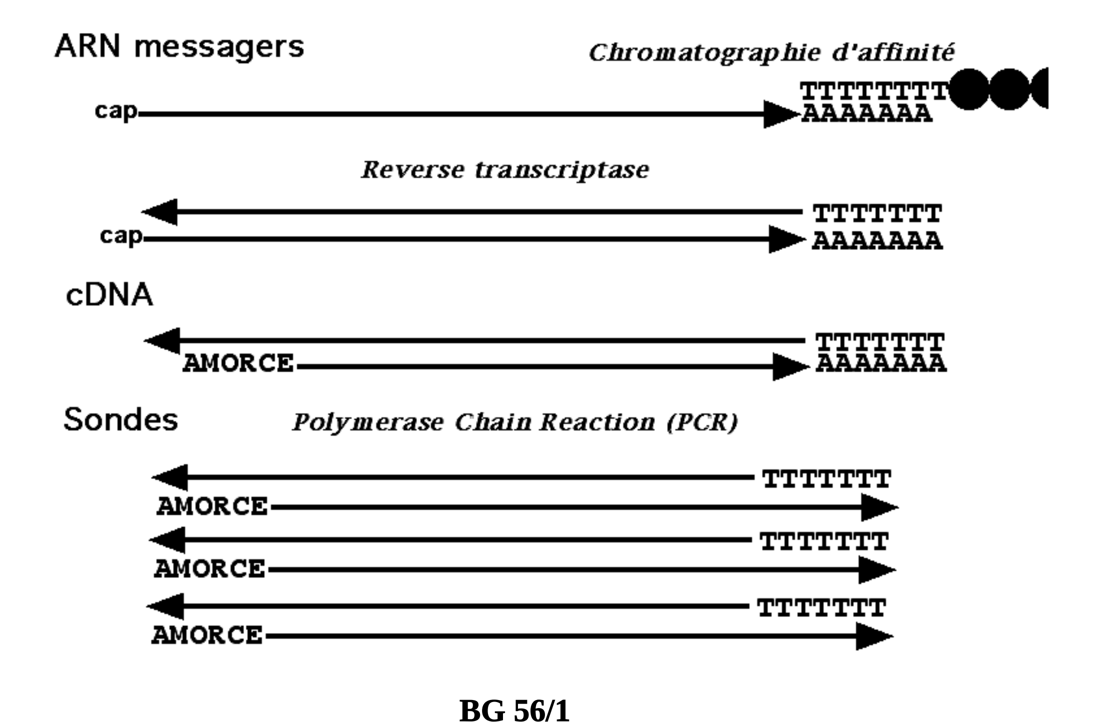


BG 56

- La manipulation et l'étude des RNA est difficile à cause de leur grande sensibilité aux ribonucléases qui les détruisent.
- Il est nécessaire de recopier la séquence en DNA pour qu'elle soit plus stable et qu'on puisse l'amplifier en fonction des besoins.
- Le mRNA purifié (chromatographie d'affinité sur une colonne d'oligo-d(T)) est lié dans un premier temps avec des oligonucléotides poly(T) qui s'attachent à la queue poly(A).
- A partir de l'extrémité 3' de cette amorce poly(T) la transcriptase réverse, qui est une DNA polymérase, synthétise un brin de DNA complémentaire du messager de départ.
- Une fois cette synthèse achevée on dégrade le RNA par une base forte ou par une ribonucléase spécifique.
- Le brin de DNA fabriqué forme spontanément à son extrémité 3' une boucle en épingle à cheveux en s'hybridant sur lui-même.
- L'extrémité 3' de cette boucle va servir de site de démarrage pour la DNA polymérase qui va synthétiser un brin de DNA complémentaire du premier.
- Une nucléase spécifique du DNA simple brin supprimera la boucle de l'extrémité. Le cDNA double brin est prêt.
- On peut ainsi constituer autant de cDNA qu'il existe de messagers dans une cellule : l'ensemble de ces cDNA forme une « banque de cDNA ».



8.19 Sondes de cDNA amplifié

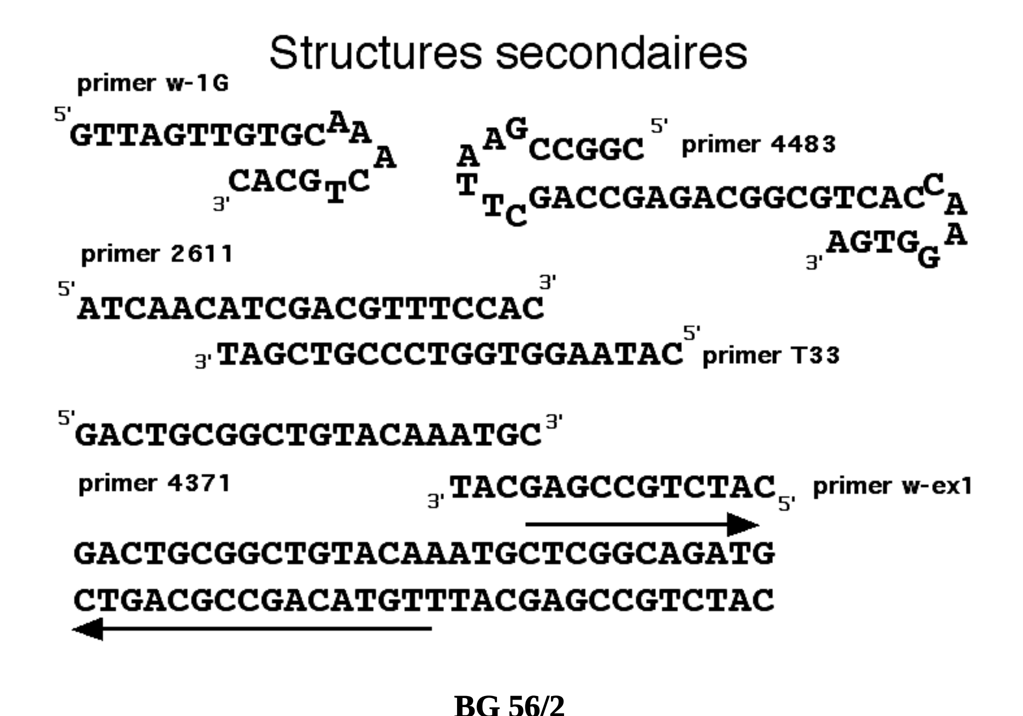


BG 56/1

- On peut effectuer directement la synthèse biologique d'une sonde à partir d'un ARN messager, isolé dans la fraction des RNA-poly(A).
- Dans une première étape, une amorce oligo d(T) est hybridée avec les séquences poly(A) des ARN messagers et la transcriptase réverse permet la synthèse d'un cDNA.
- Dans la seconde étape, une amorce spécifique permet de synthétiser le second brin d'un cDNA correspondant à l'amorce choisie.
- Enfin, la PCR permettra d'amplifier les deux brins d'ADN jusqu'à l'obtention d'une quantité suffisante de la sonde. Au cours de cette PCR on peut marquer la sonde avec un nucléotide radioactif ou biotinylé. Si la spécificité de l'amorce n'est pas suffisante, une seconde paire d'amorces spécifiques permettra d'obtenir la séquence recherchée (*nested-PCR*).



8.20 Structures secondaires

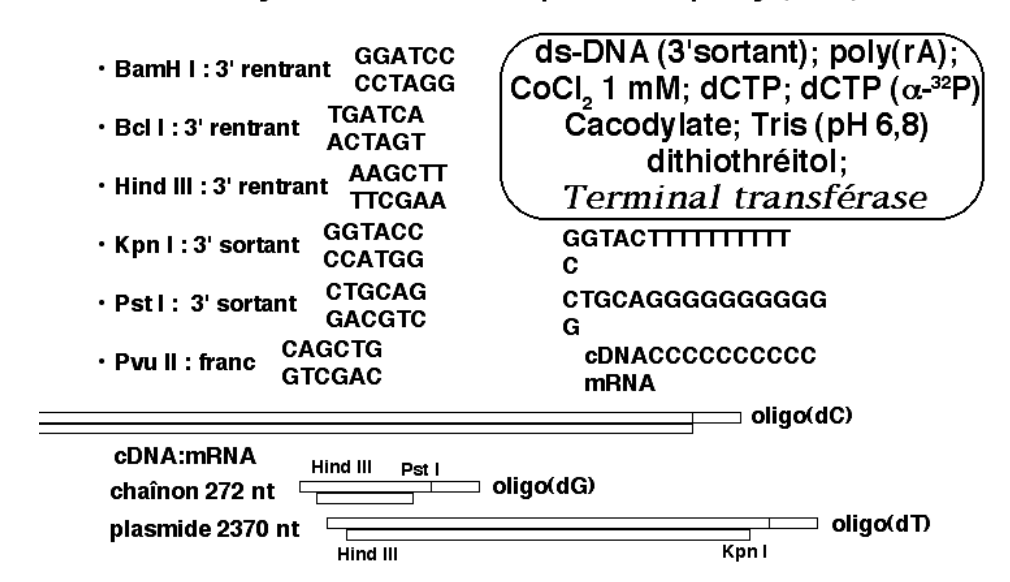


BG 56/2

- Les amorces synthétisées présentent parfois dans leur structure primaire des zones complémentaires qui s'hybrident en formant des boucles ou épingles à cheveux (*hair pin loops*).
- Lorsque ces structures s'achèvent par une extrémité 3'OH rentrante elles forment des sites de démarrage pour la polymérase. L'enzyme synthétise alors des fragments parasites à partir des amorces qui peuvent aller jusqu'à une taille double de celle des amorces hybridées.
- Des logiciels permettent de prévoir ces structures pour les éviter.



8.21 Synthèse des queues poly(dN)



BG 57

- La synthèse d'une queue poly(dN) est catalysée par la transférase terminale en présence de cations divalents : Mg^{++}, Mn^{++}, Co^{++}
- L'ADN à allonger doit présenter une extrémité 3' sortante ou à défaut un bout franc. C'est pourquoi il est utile de le digérer d'abord avec une enzyme de restriction comme Kpn I ou Pst I (bouts 3' sortants) ou à défaut comme Pvu II (bouts francs).
- L'ion Co^{++} permet la condensation préférentielle du dCTP, donc la synthèse de queues poly(dC).



Chapitre 9

Caractérisation des acides nucléiques



9.1 Principaux oligonucléotides

Oligonucléotides

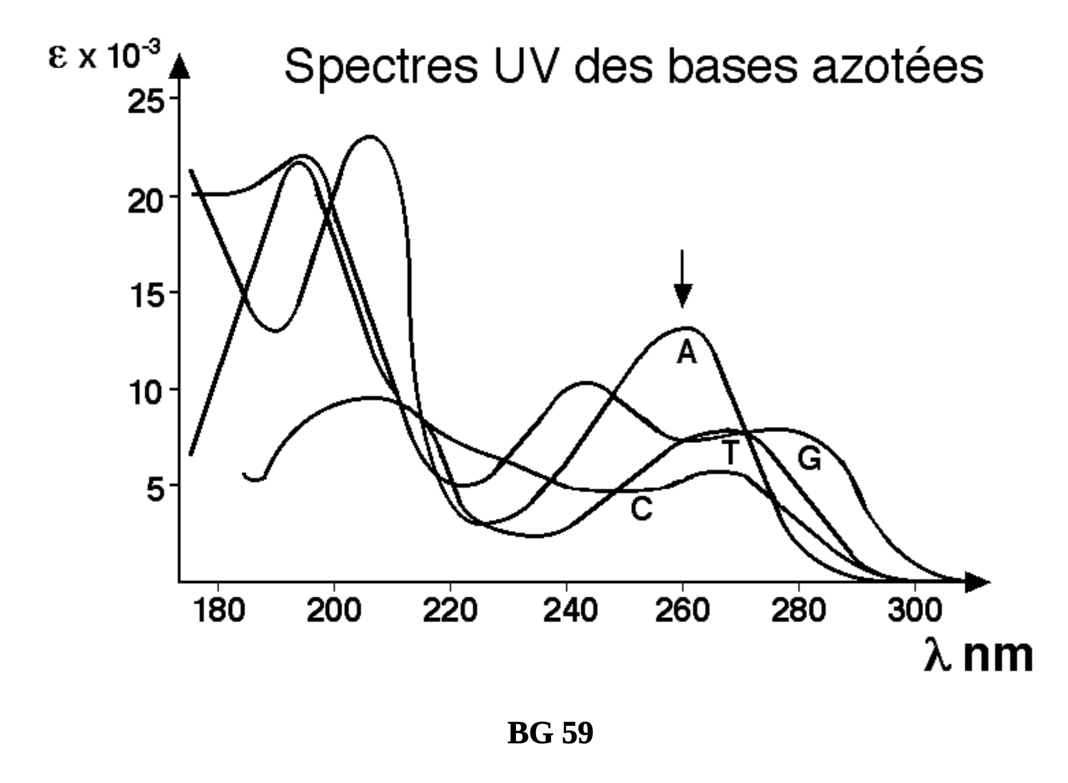
- **Amorces**
  - aléatoires 6 pb 4096 possibles
  - PCR 20 à 30 pb Hybridation en 3'
  - modifiées nouveau site de restriction mutagénèse dirigée
- **Sondes**
  - Southern 30 à 5000 pb Hybridation des exons

BG 58

- On synthétise des oligonucléotides pour de multiples techniques :
  - — des petits oligonucléotides de 6 nt pour le *random priming* ;
  - — des amorces de 20 à 30 nt pour la PCR ;
  - — des amorces modifiées pour la création des sites de restriction ou pour la mutagénèse dirigée ;
  - — des sondes enfin de 30 nt à quelques kilobases pour servir de sondes dans les Southern *blots*.



9.2 Spectres UV des bases nucléiques

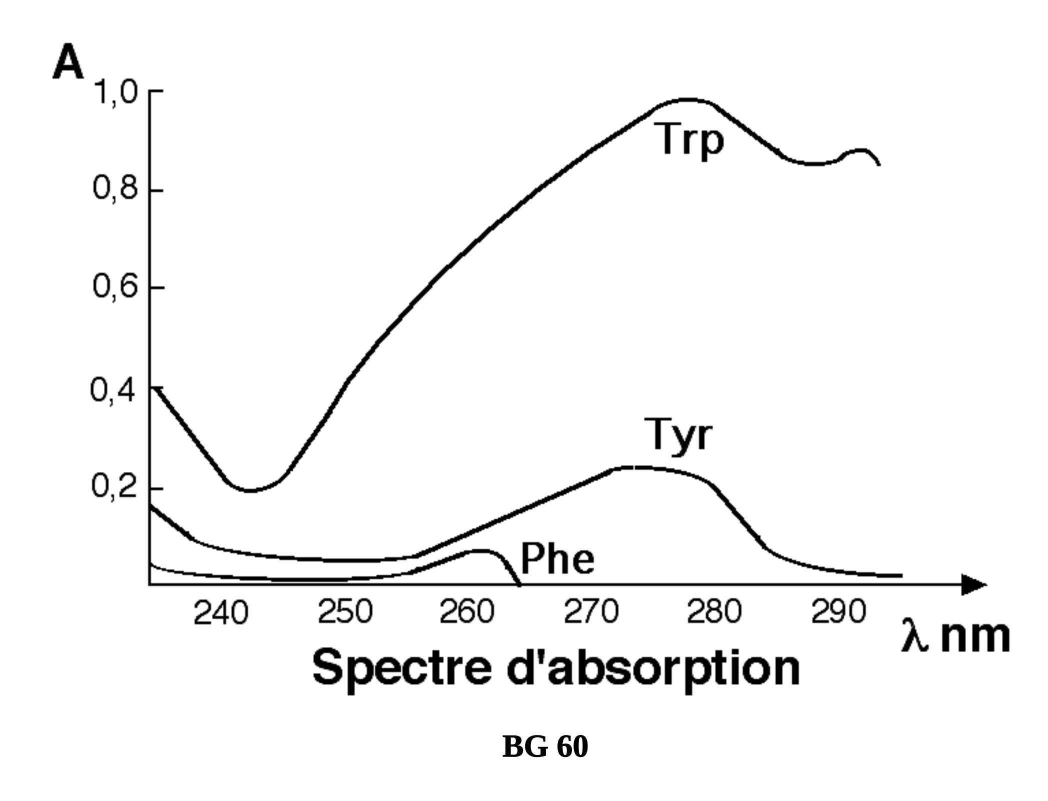


BG 59

- La plupart des molécules biologiques contenant des noyaux aromatiques absorbent les rayons ultraviolets à différentes longueurs d'onde. Dans l'ultraviolet de très courte longueur d'onde presque toutes les solutions de molécules biologiques sont opaques. Dans la région d'émission des lampes à vapeur de mercure (254 nm) les bases puriques et pyrimidiques ont des pics d'absorption spécifiques qu'on utilise pour comparer avec les spectres d'absorption des acides aminés aromatiques dans la même région.
- Sur ce graphe les ordonnées représentent l'absorption de la lumière transmise en fonction des longueurs d'onde représentées en abscisse.
- Les zones d'absorption des quatre bases s'étalent de 240 à 280 nm de sorte que les acides nucléiques formés de ces quatre types de nucléotides ont un maximum d'absorption à 260 nm.



9.3 Spectre UV des acides aminés



BG 60

- Les radicaux aromatiques de certains acides aminés (Phe, Tyr, et surtout Trp) ont la propriété d'absorber la lumière ultraviolette.
- Sur ce graphe les ordonnées représentent l'absorption de la lumière transmise en fonction des longueurs d'onde représentées en abscisse.
- L'absorption à 280 nm est due principalement aux noyaux phénols des tyrosines, parce que cet acide aminé est plus fréquent que le tryptophane, qui est pourtant beaucoup plus opaque à cette longueur d'onde.
- L'absorption de la lumière UV à 280 nm est caractéristique des protéines et sert à doser ces protéines lorsqu'elles sont en solution dans l'eau et qu'il n'existe pas d'autres molécules absorbant la lumière UV à cette longueur d'onde (acides nucléiques par exemple).



9.4 Dosage des acides nucléiques

Dosage de l'ADN

1 mole de nucléotides ≅ 309 mg/mL

Standard : 1 cm, 1 M, θ ambiante, Na^+ ≅ 0,3 M, $A_{260}/A_{280} \geq 1{,}80$

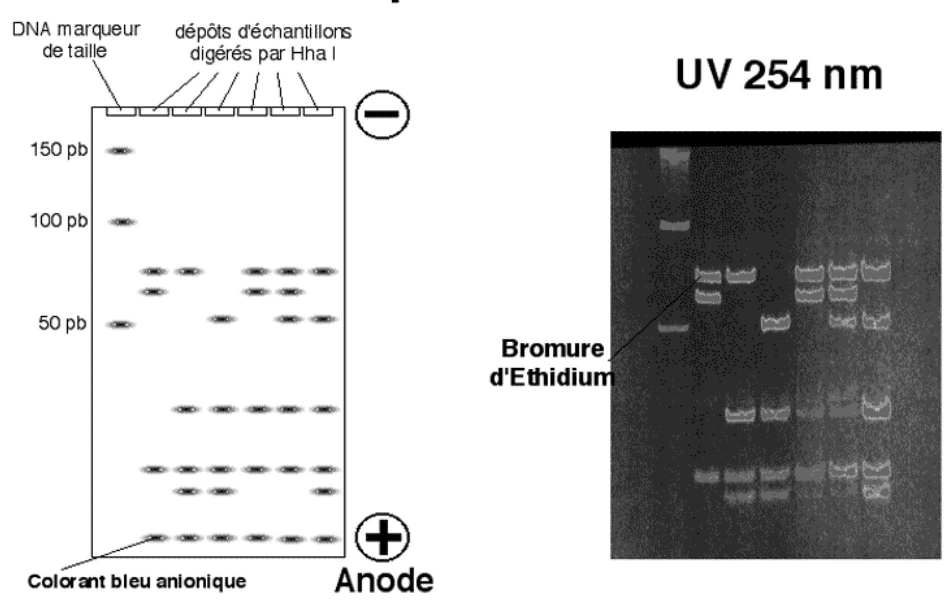
dsDNA	**$\varepsilon = 6200\ A_{260}.M^{-1}.cm^{-1}$** **$1\ A_{260} = 309/6200 = 50\ \mu g.mL^{-1}$**
ssDNA (hyperchromicité)	**$\varepsilon = 8350\ A_{260}.M^{-1}.cm^{-1}$** **$1\ A_{260} = 309/8350 = 37\ \mu g.mL^{-1}$**
RNA	**$\varepsilon = 7700\ A_{260}.M^{-1}.cm^{-1}$** **$1\ A_{260} = 309/7700 = 40\ \mu g.mL^{-1}$**

BG 61

- Pour effectuer par spectrophotométrie directe le dosage d'une solution pure d'acides nucléiques, on utilise l'absorbance des rayons ultraviolets à 260 nm. On considère dans les calculs qu'une mole de nucléotides polycondensés représente 309 grammes. Il faut se placer autant que possible dans des conditions standard de mesure : cuve de 1 cm, solution 1 M (309 mg/mL), température ambiante, $[Na^+]$ 0,3 M, rapport $A_{260}/A_{280} \geq 1{,}80$.
- Dans ces conditions le coefficient d'extinction de l'ADN double brin est de 6200 $A_{260}.M^{-1}.cm^{-1}$, c'est à dire qu'une unité d'absorbance à 260 nm (A_{260}) représente 50 µg/mL.
- Au cours de la fusion de l'ADN, le coefficient d'extinction de la solution s'élève (hyperchromicité).
- Dans ces conditions le coefficient d'extinction de l'ADN simple brin est de 8350 $A_{260}.M^{-1}.cm^{-1}$, c'est à dire qu'une unité d'absorbance à 260 nm (A_{260}) représente 37 µg/mL.
- Dans ces conditions le coefficient d'extinction de l'ARN est de 7700 $A_{260}.M^{-1}.cm^{-1}$, c'est à dire qu'une unité d'absorbance à 260 nm (A_{260}) représente 40 µg/mL.



9.5 Electrophorèse des acides nucléiques



BG 62

- Les fragments d'un DNA digéré par une enzyme de restriction comme Hha I, sont des anions et si on les soumet dans un gel à un champ électrique, ils migrent vers le pôle positif (anode) à une vitesse d'autant plus grande qu'ils sont de taille plus petite.
- On place dans un des puits de dépôt des fragments de DNA de tailles connues pour servir de marqueurs de taille.
- On ajoute dans les dépôts deux colorants :
  - — un bien visible sur le gel qui va migrer très rapidement avant les fragments de DNA pour limiter la distance parcourue au bord de l'anode ;
  - — un autre, le bromure d'éthidium qui se fixe spécifiquement au DNA quelle que soit la séquence et qui émet une fluorescence mauve lorsqu'on l'éclaire avec des rayons ultra-violets.
- On examine le gel sous lumière ultra-violette à 254 nm et on photographie les fragments d'ADN digéré.



9.6 Sonde nucléique (définition)

SONDE NUCLEIQUE :

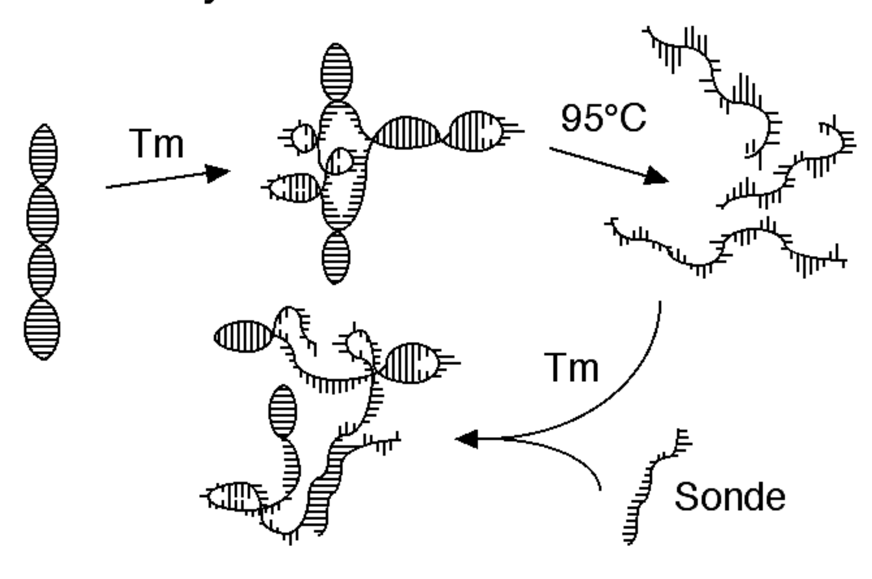
- **Molécule d'acide nucléique, antiparallèle et complémentaire d'une séquence spécifique d'un génome ou d'une banque, capable de reconnaître par hybridation et de marquer cette séquence pour permettre son identification ou son isolement.**

BG 63

- Une sonde est un marqueur qu'on introduit dans un milieu biologique pour que ses propriétés vis-à-vis des constituants de ce milieu soient révélées par l'émission d'un signal détectable par nos instruments de mesure.
- Les oligonucléotides sont employés comme sondes parce qu'ils peuvent s'hybrider spécifiquement avec une séquence d'un acide nucléique dans un milieu biologique. Lorsqu'elles sont marquées (atomes radioactifs, biotine, fluorochromes) leur présence dans le milieu peut être suivie.
- Après lavage du milieu pour éliminer les oligonucléotides non-hybridés avec le matériel biologique, la présence du marqueur révèle les hybrides formés et permet de mettre en évidence la séquence reconnue par la sonde.
- Cette technique s'applique à des coupes de tissus (hybridation *in situ*), à des clones bactériens transfectés de vecteurs recombinés (criblage des banques), etc...



9.7 Hybridation d’une sonde



BG 64

- Un fragment de DNA double brin affecte normalement une structure secondaire en double hélice.
- En élevant la température jusqu’à la Tm, on disjoint 50 % environ des liaisons hydrogène qui unissent les brins d’ADN, ce qui dénature partiellement la double hélice.
- Lorsque la température atteint 95°C. toutes les liaisons hydrogène sont rompues et la dénaturation de l’ADN est complète : ADN simple brin, qualifié de « dénaturé ». Au cours de la dénaturation, le coefficient d’extinction de l’ADN à 260 nm augmente (hyperchromicité).
- En refroidissant brutalement (glace) les structures secondaires ne se reforment pas et l’ADN reste dénaturé. Si on refroidit doucement, la double hélice se reforme progressivement. Un oligonucléotide (sonde) ajouté à ce moment peut s’hybrider avec un fragment complémentaire de l’ADN, dès que la température descend en-dessous de Tm.
- L’hybridation d’une sonde marquée (atomes radioactifs, radicaux fluorescents ou ligands spécifiques) sur un ADN dénaturé permet de marquer spécifiquement tous les fragments de cet ADN dont la séquence est complémentaire de la sonde.



9.8 Calcul de la Tm

Calcul de la Tm

- **Oligonucléotide inférieur à 20 nt**
  - (A+T)x2 + (G+C)x4 = Tm en °C.
- **Oligonucléotide supérieur à 20 nt**
  - [(A+T)x2 + (G+C)x4] x (1+[(N-20)/20]) = Tm en °C.

BG 65

- Il est possible de mesurer directement la température de fusion d'un ADN double brin en mesurant l'augmentation de l'absorbance de la solution à 260 nm en fonction de la température.
- Toutefois, on se contente la plupart du temps d'une estimation à partir de la composition de l'oligonucléotide. Si celui-ci a une longueur égale ou inférieure à 20 nt on compte 2°C par couple A:T et 4°C par couple G:C. A partir de N = 20, on corrige d'un multiplicateur proportionnel à la longueur au delà de ce chiffre : 1 + [(N-20)/20].
- Pour être plus précis il faut tenir compte aussi de la concentration en sodium du tampon d'hybridation. Lorsque cette concentration n'excède pas 1M, on utilise la formule :

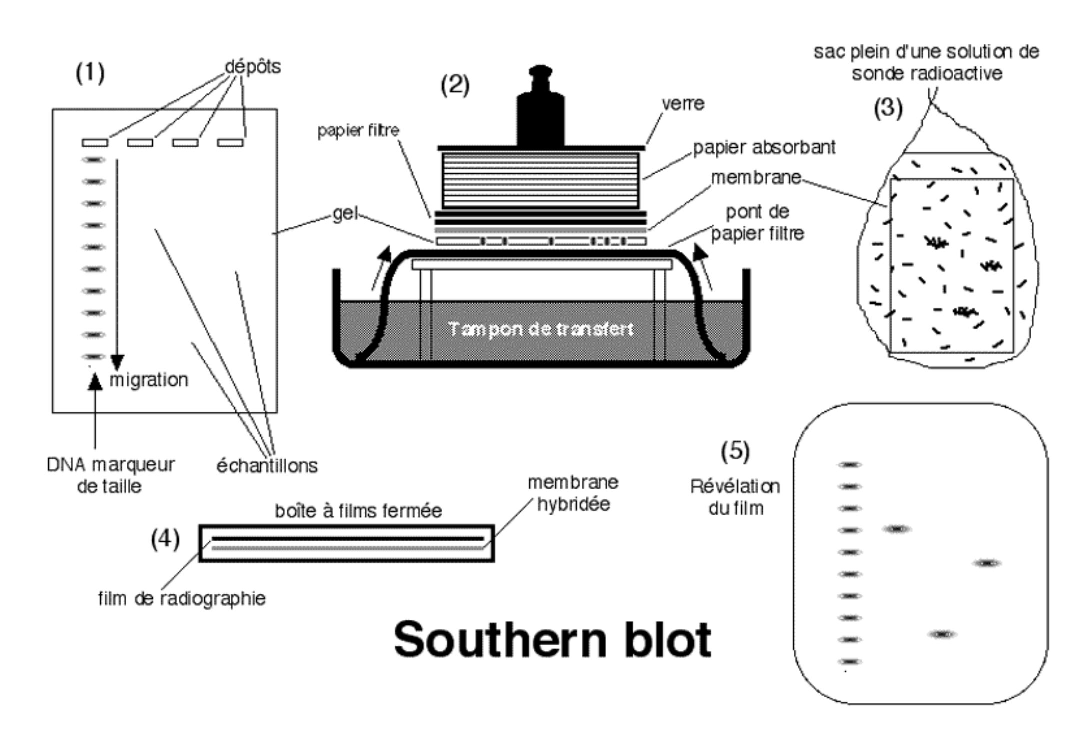
$$Tm = 81,5 + 16,6(\log_{10}[Na^{+}]) + 0,41[(G+C)/N] - (600/N)$$

- Lorsqu'il existe des misappariements (*mismatches*) il faut soustraire de la Tm calculée autant de degrés C. que le pourcentage de séquence non-appariée de l'oligonucléotide.
- Lorsqu'on travaille en présence de formamide, il faut encore diminuer la Tm en fonction de la concentration de l'amide :

$$T'm = Tm-0,6(\% \text{ formamide})$$



9.9 Southern blot



BG 66

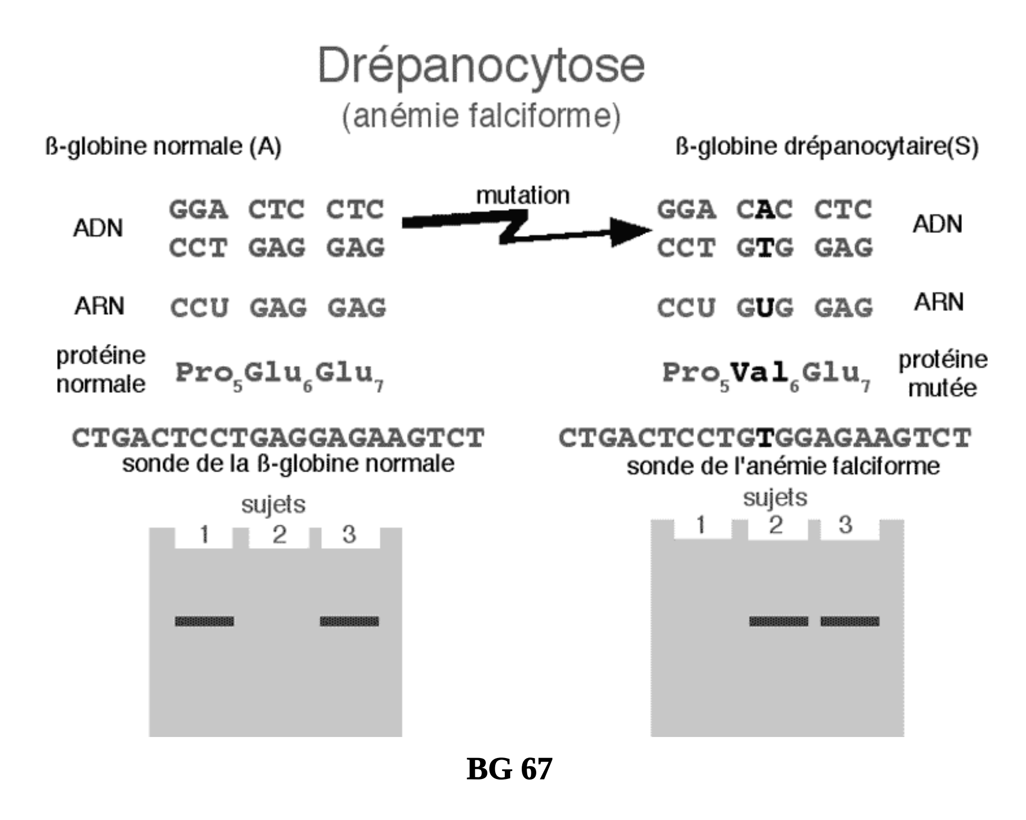
- Le Southern blot est une technique mise au point par Edward M. Southern en 1975 pour rechercher des fragments de DNA sur une électrophorèse en les hybridant avec une sonde complémentaire.
    1. Après avoir digéré le DNA par une enzyme de restriction, on obtient un mélange de très nombreux fragments de restriction. On soumet ces fragments à un électrophorèse pour les faire migrer dans un gel de haut en bas en fonction inverse de leur taille.
    2. On fait un montage pour faire passer les fragments grâce à une montée de tampon imprégnant le gel puis une membrane de nylon où le DNA va se fixer par des liaisons stables.
    3. La membrane de nylon avec le DNA fixé est alors mise à incuber dans un sac contenant une solution d’une sonde radioactive complémentaire du fragment de DNA qu’on recherche, à une température assez basse pour que l’hybride se forme mais assez élevée pour que cet hybride soit parfaitement complémentaire.
    4. On lave la membrane des molécules de la sonde qui ne sont pas fixées à leur DNA complémentaire, puis on la met en présence d’un film radiographique vierge dans une enceinte opaque pour que la sonde radioactive fixée sur les fragments de DNA complémentaires impressionne le film.
    5. On révèle le film où des taches noires (sur le négatif) correspondent aux emplacements où ont migré les fragments d’ADN complémentaires de la sonde. On compare les distances de migration avec des fragments de DNA radioactifs de tailles connues qui servent



de marqueurs de taille.



9.10 Drépanocytose (anémie falciforme)

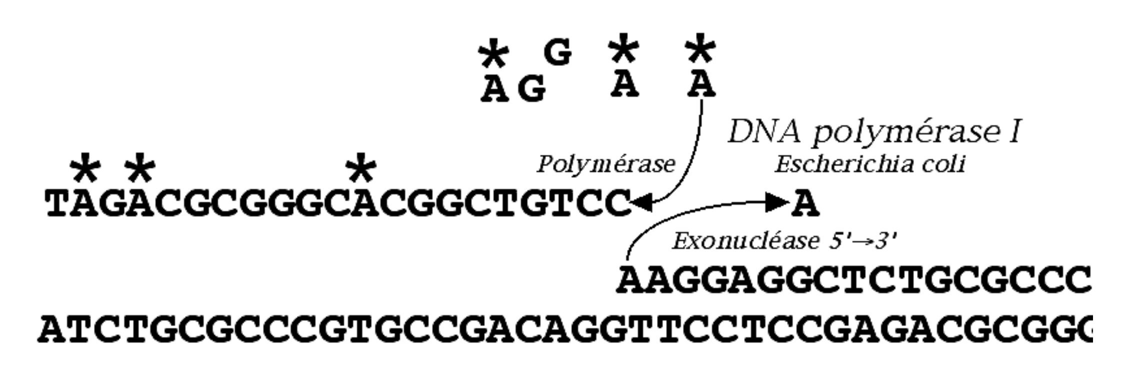


BG 67

- La drépanocytose ou anémie falciforme est une maladie des globules rouges qui sont déformés par une cristallisation anormale de l'hémoglobine. Cette cristallisation résulte d'une substitution A→T dans le sixième codon.
- Cette substitution s'exprime par une mutation Glu6→Val de la chaîne β de la globine.
- L'électrophorèse de l'ADN du gène HBB, qui code pour la chaîne ß des globines est révélée par deux sondes : l'une spécifique de la séquence normale, l'autre de celle de la protéine mutée.
- L'ADN d'un sujet est révélé par la seule sonde normale s'il possède deux gènes HBB normaux (homozygote normal), par la seule sonde de la protéine mutée s'il possède deux gènes HBB responsables de la mutation (homozygote muté) et par les deux sondes s'il possède un gène normal et un gène de la protéine mutée (hétérozygote).
- Les sujets homozygotes mutés sont atteints d'anémie falciforme et sujets à des crises graves. Les sujets hétérozygotes sont porteurs du « trait drépanocytaire », ils n'ont que peu de troubles mais ils transmettent le gène à leurs descendants.



9.11 Marquage : *nick translation*



BG 68

- Le « déplacement de brèche » (*nick translation*) est une technique de marquage utilisant les propriétés de la DNA polymerase I (*E. coli*), pour obtenir des sondes radioactives.
- L'ouverture d'une brèche dans la structure primaire d'un des deux brins d'un fragment d'ADN est obtenue par l'action de la désoxyribonucléase pancréatique.
- La DNA polymérase synthétise un nouveau brin d'ADN à partir de l'extrémité 3'OH terminale d'un coté de la brèche ainsi créée. Simultanément l'activité 5'→3' exonucléase de la DNA pol I hydrolyse les nucléotides du coté 5'-phosphate de la même brèche. Il en résulte un déplacement de la brèche le long de l'ADN avec un remplacement des nucléotides du brin coupé.
- Si cette opération est réalisée en présence de nucléotides substrats fortement marqués, le marqueur sera incorporé dans le DNA synthétisé. Pour éviter la formation de structures en épingles à cheveux, la synthèse doit être maintenue très lente à une température basse : 16° C



9.12 Marquage : *random priming*

Random priming
(Amorçage aléatoire)

CGTACA
GTACGG
GTTAAG

Nucléotides marqués
*A *A G *A C G T

Hexanucleotide
amorce

DNA polymérase I
Escherichia coli
(fragment de Klenow)

Polymérase

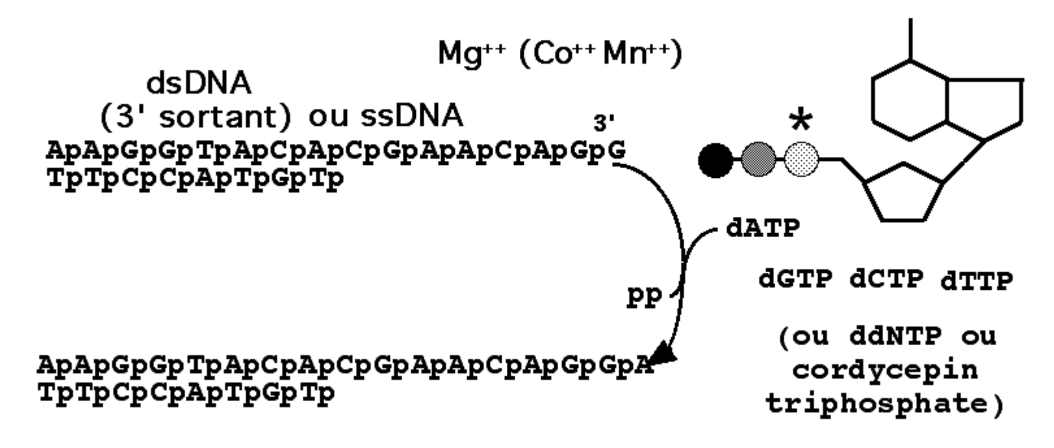
CGGCACGGCTGTCC
ATCTGCGCCCGTGCCGACAGGTTCCTCCGAGACGCGGC

BG 69

- La synthèse d’un brin marqué de DNA peut aussi être faite par une DNA polymérase à partir d’amorces contenant des séquences aléatoires. Ces amorces sont obtenues par synthèse chimique d’oligonucléotides contenant au hasard chacune des quatre bases à chaque position. On peut aussi utiliser un produit de digestion poussée d’un ADN génomique réduit en fragments aléatoires de 6 à 12 nucléotides.
- Dans les deux cas, un ou plusieurs des oligonucléotides s’hybrident avec la séquence en 5’ du fragment d’intérêt (ADN simple brin) et servent d’amorces pour la synthèse d’un deuxième brin.
- Si cette opération est réalisée en présence de nucléotides substrats fortement marqués, le marqueur sera incorporé dans le DNA synthétisé.



9.13 Marquage : terminal transferase

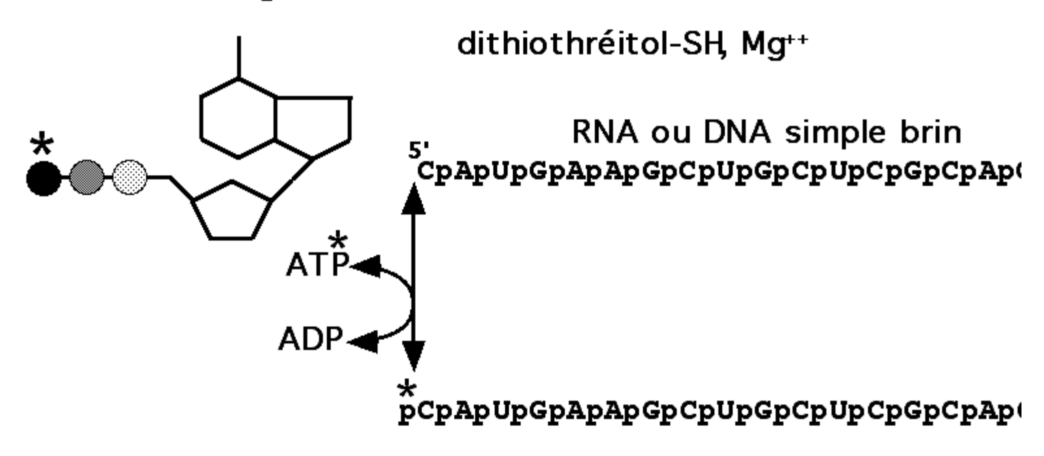


BG 70

- La transférase terminale catalyse l'addition de désoxynucléotides sur une fonction alcool 3'OH terminale de l'ADN. L'enzyme préfère les molécules d'ADN dont l'extrémité 3'OH est sortante, mais il existe des conditions qui favorisent la catalyse sur l'ADN à bouts francs, voire sur les extrémités 3'OH rentrantes. Elle incorpore plus spécifiquement les purines ou les pyrimidines en fonction du cation divalent qu'on lui donne comme cofacteur : Mg^{++} pour les purines, Co^{++} pour les pyrimidines ou encore Mn^{++}.
- La transférase terminale du commerce est extraite du thymus de veau. Une unité d'enzyme catalyse le transfert de 17 picomoles de désoxyribonucléotide par minute à 37° C.
- La transférase terminale est utilisée au laboratoire pour :
  - construire des séquences polymérisées de nucléotides (queues) à l'extrémité 3'OH de l'ADN (en vue du clonage des fragments d'ADN) ;
  - marquer les extrémités 3' de l'ADN avec un nucléotide marqué sur le phosphate α ;
  - ajouter un nucléotide à la fin d'une séquence pour induire une mutation (mutagénèse dirigée).



9.14 Marquage : polynucléotide kinase



BG 71

- La polynucléotide kinase catalyse le transfert du phosphate γ de l'ATP sur une fonction alcool du carbone 5' d'un ADN ou d'un ARN. Le substrat requiert une déphosphorylation préalable si la fonction est estérifiée. L'enzyme catalyse aussi l'échange du phosphate 5' terminal d'un ADN ou d'un ARN avec le phosphate γ de l'ATP, en présence d'ADP comme accepteur de phosphate.
- La polynucléotide kinase du commerce est extraite d'une bactérie (*E. coli*) infectée par le bactériophage T4. Une unité d'enzyme catalyse le transfert de 33 picomoles de phosphate par minute à 37° C.
- La polynucléotide kinase est utilisée au laboratoire pour incorporer du phosphate radioactif (^{32}P) sur l'extrémité 5' d'un acide nucléique, soit par transfert, soit par échange.



9.15 Didésoxyadénosine triphosphate

didésoxyadénosine triphosphate

Adénine

3 phosphates

didésoxy-β-D-Ribose

BG 71/5

- Le didésoxyadénosine triphosphate (ddA) est un nucléotide composé de synthèse. Sa structure est dépourvue de fonction alcool en 3’.
- Analogue structural de nucléotide, le ddA est utilisé comme inhibiteur de la réplication (DNA polymérase). L’absence de fonction alccol en 3’ empêche toute condensation avec le nucléotide suivant.
- Cette inhibition est principalement utilisée pour le séquençage des DNA.



9.16 Réaction de séquence

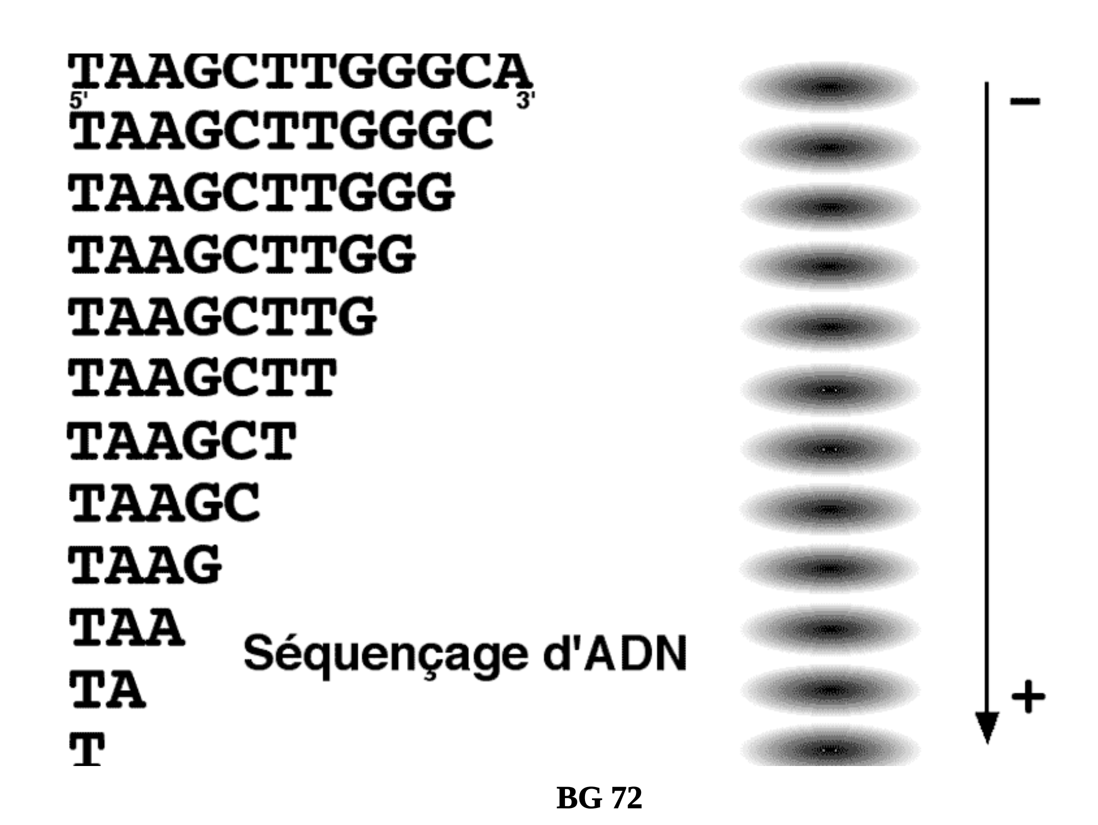
Extrémité 3'OH du DNA synthétisé
Liaison ester nouvelle
Condensation
H_2O
JOE
Hydrolyse des liaisons riches en énergie
Pas de 3'-OH
ddATP

BG 71/6

- Pour lire la séquence d'un DNA simple brin on hybride une amorce du côté 3' de ce DNA. Puis on effectue avec une DNA polymérase la synthèse d'un brin complémentaire.
- Pour cette synthèse on apporte comme substrats les quatre désoxyribonucléosides triphosphates (dATP, dCTP, dGTP et dTTP). En plus une très petite quantité de didésoxyribonucléosides triphosphates fluorescents (ddATP-JOE, ddCTP-5-FAM, ddGTP-TAMRA et ddTTP-ROX). La polymérase choisira le plus souvent un désoxyribonucléoside normal et la synthèse se poursuivra jusqu'à ce qu'elle incorpore un didésoxyribonucléoside fluorescent.
- A ce stade le brin en cours de synthèse n'a plus d'extrémité 3'-OH et la réaction de polycondensation ne peut plus se poursuivre.
- Le didésoxyribonucléotide incorporé en dernier est fluorescent et émet sous l'excitation une lumière verte pour JOE (ddAMP), bleue pour 5-FAM (ddCMP), jaune pour TAMRA (ddGMP) et rouge pour ROX (ddTMP).
- A chaque lettre de la polycondensation un petit nombre de molécules sont ainsi arrêtées et marquées de la fuorescence correspondant au dernier nucléotide incorporé.
- En séparant ces molécules par électrophorèse en fonction de leur taille on peut lire les lettres successives qui apparaissent comme des zones sur l'électrophorégramme dont la fuorescence correspond à la base de ce dernier nucléotide.



9.17 Séquençage de l'ADN



BG 72

- Pour connaître la séquence des ADN, on fait synthétiser un brin d'ADN par une enzyme spécifique.
- L'enzyme commence son travail à partir de l'extrémité 3' d'une sonde hybridée qui sert d'amorce. Elle ajoute des nucléotides complémentaires de ceux du brin d'ADN qu'elle copie.
- On lui donne pour substrats des désoxynucléotides triphosphates normaux mélangés avec des didésoxynucléotides dont la fonction alcool secondaire en 3' est réduite ce qui empêche la synthèse de se poursuivre au delà.
- Les didésoxynucléotides incorporés en dernier sont marqués spécifiquement par des molécules fluorescentes (vert pour didésoxyA, bleu pour didésoxyC, jaune pour didésoxyG et rouge pour didésoxyT).
- On sépare ensuite les fragments synthétisés dans un champ électrique (électrophorèse : les ADN sont des anions, ils vont donc vers le pôle +), en fonction de leur longueur (les plus petits vont plus vite).
- On lit ensuite les taches successives, identifiées par leur couleur, ce qui révèle la séquence des fragments synthétisés.



9.18 Séquençage : *dye primers*

Dye primers

```
JOE   5' TGTAAAACGACGGCCAGTGCGGGCACGGCTGTCCAA 3' + ddATP
ROX   5' TGTAAAACGACGGCCAGTGCGGGCACGGCTGTCCAA 3' + ddTTP
5-FAM 5' TGTAAAACGACGGCCAGTGCGGGCACGGCTGTCCAA 3' + ddCTP
TAMRA 5' TGTAAAACGACGGCCAGTGCGGGCACGGCTGTCCAA 3' + ddGTP
                            3'ATTCGAACCGTGCCGACAGGTTCCTCGACGTCCGCCGCGTCCGGGC
                            5'TAAGCTTGGCACGGCTGTCCAAGGAGCTGCAGGCGGCGCAGGCCCG

CGACCCGCGCCTGTACCTCCTGCACGCGCCGGCGGACCACGTCATGGCGCCGCTCCACGTCCGGTACGAGCC
GCTGGGCGCGGACATGGAGGACGTGCGCGGCCGCCTGGTGCAGTACCGCGGCGAGGTGCAGGCCATGCTCGG

GGTCTCGTGGCTCCTCGACGGGGTGCGCCTCGCCTCCCACCTGCGCAAGCTGCGTAAGCGGCTCCTCCGCGA
CCAGAGCACCGAGGAGCTGCCCCACGCGGAGCGGAGGGTGGACGCGTTCGACGCATTCGCCGAGGAGGCGCT

TGCCGATGACCTGCAGATCGCGGACCGTCACATGGTCCGGCCCCGCTTAAGACA 5'
ACGGCTACTGGACGTCTAGCGCCTGGCAGTGTACCAGGCCGGGGCGAATTCTGT 3'
          ddCTP + 3' CGTCACATGGTCCGGCCCCCAGTATCGACAAAGGAC 5' 5-FAM
          ddTTP + 3' CGTCACATGGTCCGGCCCCCAGTATCGACAAAGGAC 5 ROX
          ddATP + 3' CGTCACATGGTCCGGCCCCCAGTATCGACAAAGGAC 5 JOE
          ddGTP + 3' CGTCACATGGTCCGGCCCCCAGTATCGACAAAGGAC 5 TAMRA
```

BG 73

- La réaction de séquençage est une technique de synthèse de DNA analogue à la PCR. On utilise des amorces marquées (*dye primers*) ou des nucléotides marqués (*dye terminators*).
- Pour séquencer ce fragment de 244 nt on utilisera quatre amorces différentes pour chacun des deux brins d'ADN :
  - — la première amorce liée au fluorochrome JOE (vert) se termine par une séquence complémentaire de l'extrémité 3' du brin du dessus ; elle sera incubée avec le tampon, la DNA polymérase, les quatre dNTP et en addition une quantité plus faible du didésoxynucléotide ddATP (nucléotide dont les carbones 2 et 3 du ribose n'ont plus de fonctions alcool). L'incorporation de ce dernier va interrompre la synthèse de façon aléatoire mais toujours au niveau d'un A de la séquence.
  - — la deuxième amorce marquée ROX (rouge) sera arrêtée par du ddTTP
  - — la troisième amorce marquée 5-FAM (bleu) sera arrêtée par du ddCTP
  - — la quatrième amorce enfin, marquée par TAMRA (jaune) sera arrêtée par du ddGTP.
- Tous les fragments synthétisés seront mis ensemble dans une électrophorèse qui les séparera en fonction de leur masse moléculaire donc du nombre de nucléotides et on enregistrera la couleur de l'amorce qui indiquera spécifiquement les fragments se terminant par chacun des quatre didésoxyribonucléotides. La séquence de ces couleurs indique la séquence de l'ADN : vert pour A, rouge pour T, bleu pour C et jaune pour G.
- On contrôle enfin la séquence obtenue par celle obtenue avec des amorces s'hybridant avec



le brin du dessous, cette seconde séquence devant être antiparallèle et complémentaire de la première.



9.19 Séquençage : *dye terminators*

Dye terminators

ddCTP-5-FAM
ddTTP-ROX
ddGTP-TAMRA

dNTP +ddATP-JOE

```
                            5'TAAGCTTGGCACGGCTGTCCAA 3'
                            3'ATTCGAACCGTGCCGACAGGTTCCTCGACGTCCGCCGCGTCCGGGC
                            5'TAAGCTTGGCACGGCTGTCCAAGGAGCTGCAGGCGGCGCAGGCCCG

CGACCCGCGCCTGTACCTCCTGCACGCGCCGGCGGACCACGTCATGGCGCCGCTCCACGTCCGGTACGAGCC
GCTGGGCGCGGACATGGAGGACGTGCGCGGCCGCCTGGTGCAGTACCGCGGCGAGGTGCAGGCCATGCTCGG

GGTCTCGTGGCTCCTCGACGGGGTGCGCCTCGCCTCCCACCTGCGCAAGCTGCGTAAGCGGCTCCTCCGCGA
CCAGAGCACCGAGGAGCTGCCCCACGCGGAGCGGAGGGTGGACGCGTTCGACGCATTCGCCGAGGAGGCGCT

TGCCGATGACCTGCAGATCGCGGACCGTCACATGGTCCGGCCCCGCTTAAGACA 5'
ACGGCTACTGGACGTCTAGCGCCTGGCAGTGTACCAGGCCGGGGCGAATTCTGT 3'
                            3'CGTCACATGGTCCGGCCCCCTTAAGACA 5'
```

dNTP +ddATP-JOE

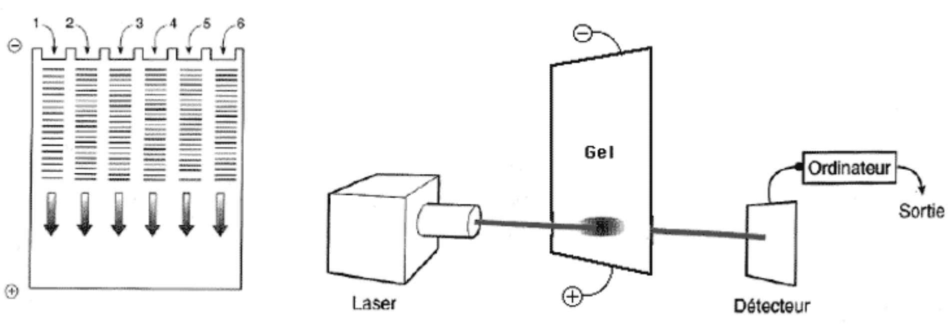
BG 73/1

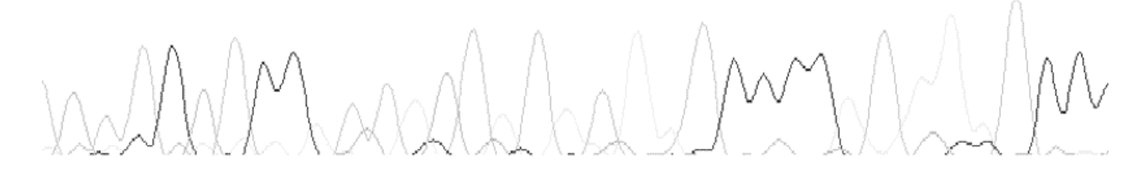
- La réaction de séquençage est une technique de synthèse de DNA analogue à la PCR. On utilise des amorces marquées (*dye primers*) ou des nucléotides marqués (*dye terminators*).
- Pour séquencer ce fragment de 244 nt on utilisera une seule amorce spécifique pour chacun des deux brins d'ADN et complémentaire de l'extrémité 3' de ce brin. Elle sera incubée avec le tampon, la DNA polymérase, les quatre dNTP et en addition une quantité plus faible des quatre didésoxynucléotides ddATP (lié au fluorochrome JOE = vert), ddCTP (lié au fluorochrome 5-FAM = bleu), ddGTP (lié au fluorochrome TAMRA = jaune) et ddTTP (lié au fluorochrome ROX = rouge). L'incorporation de ces didésoxynucléotides marqués va interrompre la synthèse de façon aléatoire mais toujours au niveau de la lettre correspondante de la séquence.
- Tous les fragments synthétisés seront mis dans une électrophorèse qui les séparera en fonction de leur masse moléculaire donc du nombre de nucléotides et on enregistrera la couleur du fluorochrome qui indiquera spécifiquement les fragments se terminant par chacun des quatre didésoxyribonucléotides marqués. La séquence de ces couleurs indique la séquence de l'ADN : vert pour A, rouge pour T, bleu pour C et jaune pour G.
- On contrôle enfin la séquence obtenue par celle obtenue avec une amorce s'hybridant avec le brin du dessous, cette seconde séquence devant être antiparallèle et complémentaire de la première.



9.20 Gel de séquence

Gel de séquence



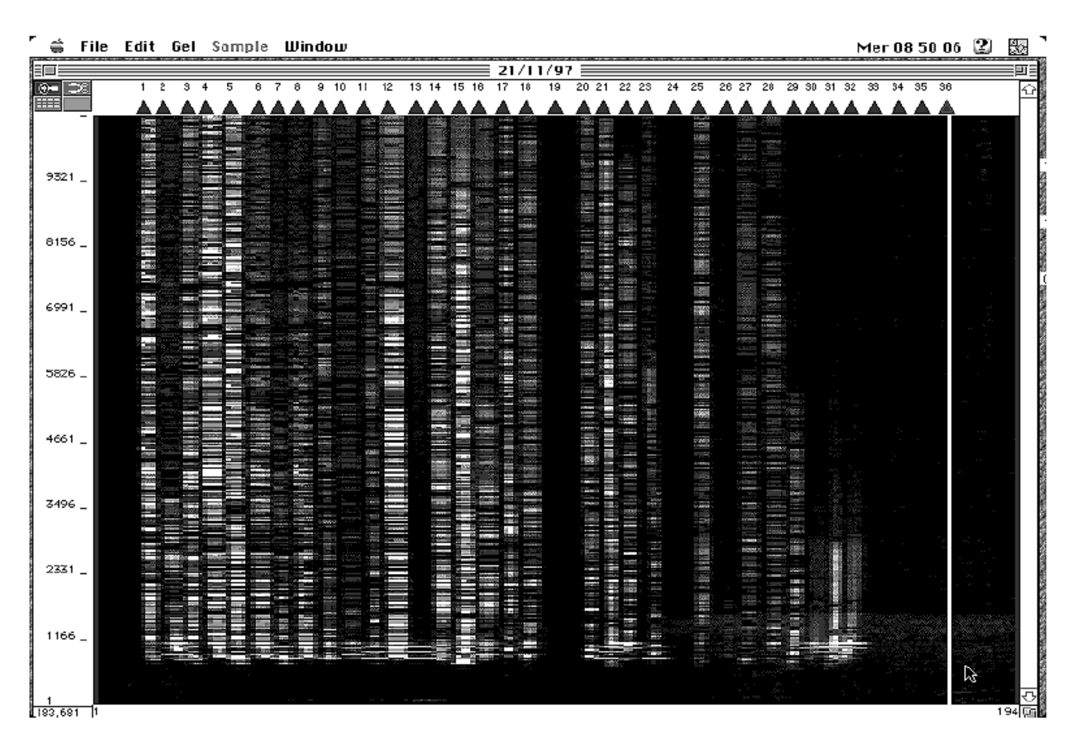


BG 73/2

- En utilisant des amorces spécifiques, on fait synthétiser un brin complémentaire du DNA à lire, en présence d'une faible proportion de didésoxyribonucléotides marqués d'un radical fluorescent différent pour chaque base.
- La DNA polymérase lorsqu'elle a incorporé un tel didésoxyribonucléotide ne peut plus continuer la synthèse faute d'extrémité 3'OH libre. Il se forme donc une multitude de brin complémentaires inachevés, chacun terminé par un didésoxyribonucléotide fluorescent caractéristique de la base de ce dernier nucléotide.
- En séparant par électrophorèse ces fragments complémentaires on sépare dans le gel chacun des fragments, du plus petit au plus grand et on les détecte au passage par un faisceau laser qui excite la fluorescence et une cellule photoélectrique qui lit la lumière émise à chacune des longueurs d'onde des fluorescences caractéristiques des quatre bases.
- L'ordinateur reçoit donc une série de mesure d'intensité lumineuses en forme de pics correspondant au passage de chacun des fragments : en interprétant la couleur de la fluorescence de chaque pic l'ordinateur écrit la séquence de l'ADN complémentaire.



9.21 Séquençage : image du gel

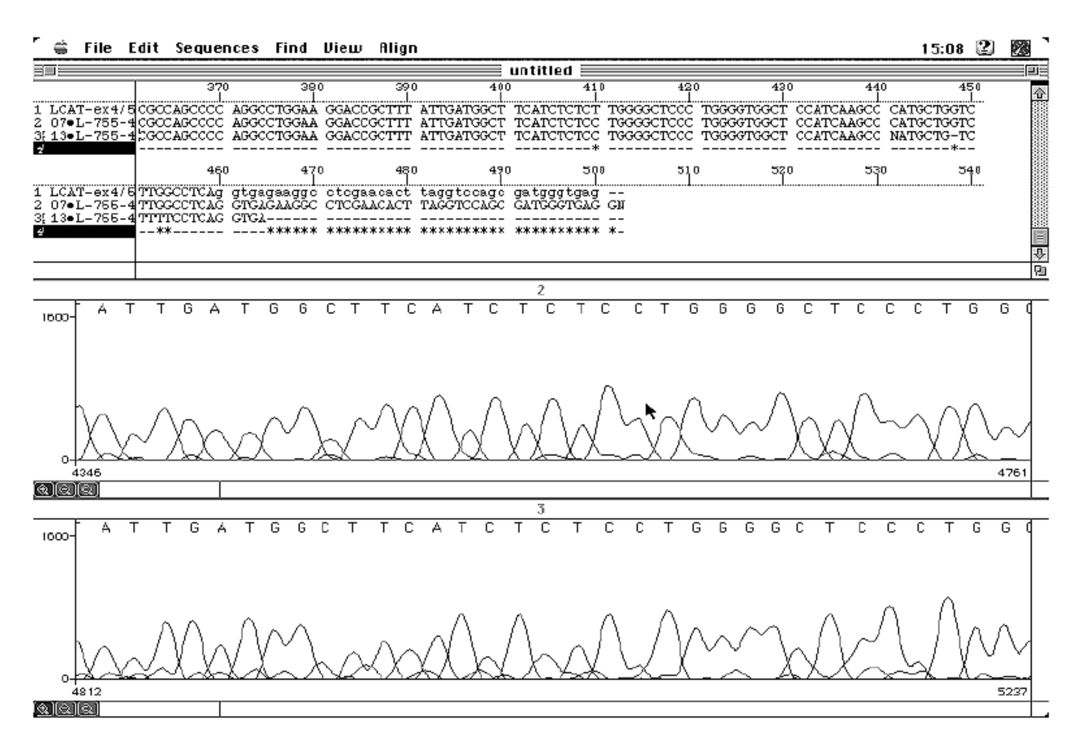


BG 74

- Le gel de séquençage est un gel de polyacrylamide très fin dans lequel migrent des fragments d'acides nucléiques incolores, marqués à leur extrémité 3' par un didésoxynucléotide fluorescent (ddN*). Le laser émet une lumière d'excitation qui passe à travers le gel et les ddN* excités réémettent une lumière d'une longueur d'onde plus grande spécifique de chaque ddN*, qui sera lue par une caméra à quatre longueurs d'onde différentes, chacune spécifique d'un des quatre didésoxynucléotides. La caméra enregistre donc en fonction du temps l'intensité de la lumière émise par le gel dans les quatre longueurs d'onde.
- Afin de permettre de suivre la progression de l'électrophorèse sur l'écran, l'ordinateur construit une image virtuelle du gel : le gel y est représenté avec des taches de couleurs différentes pour les quatre ddN* (vert pour A, jaune pour G, bleu pour C et rouge pour T) et d'intensité proportionnelle à la lecture du gel. Cette image virtuelle est assez longue pour représenter toute l'électrophorèse ; les fragments les plus rapides qui sont depuis longtemps passés dans la cuve de l'anode, sont toujours visibles en haut de l'image pendant que le bas de l'image représente les fragments qui parviennent réellement au même moment à l'endroit du gel où se fait la lecture.
- Cette image permet de guider (*tracking*) la lecture du tracé de chaque puits de dépôt, même si celui-ci n'a pas migré dans le gel en parfaite ligne droite.



9.22 Séquençage : analyse de l’image

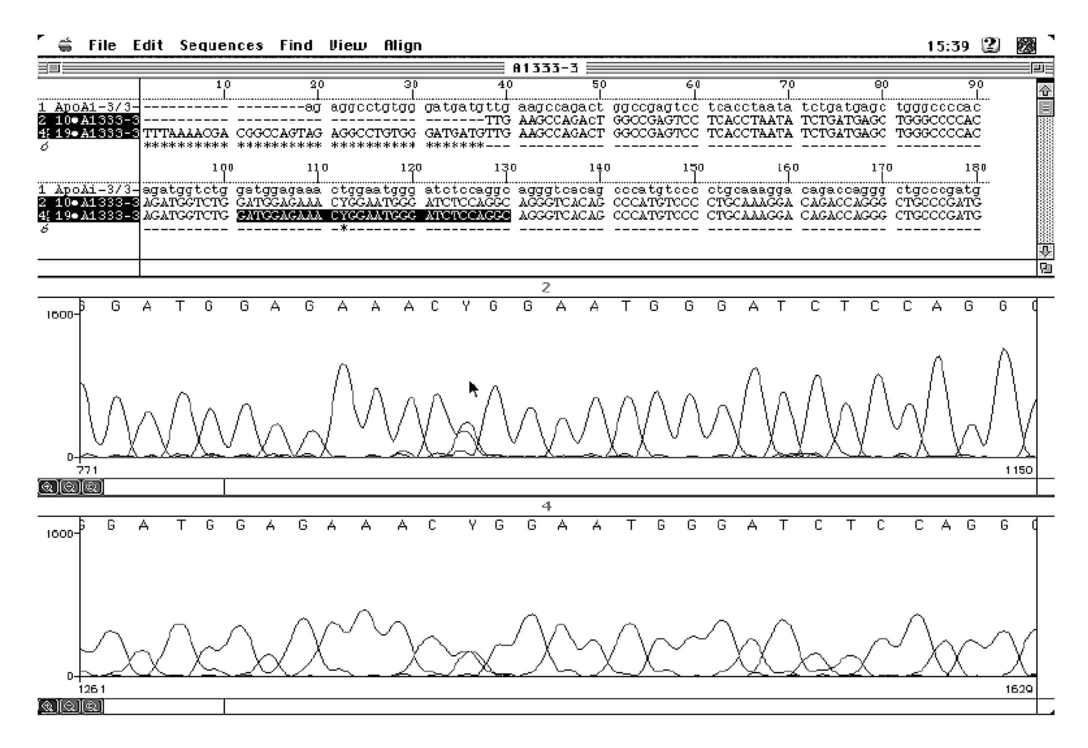


BG 74/1

- La lecture des fluorescences permet de tracer quatre courbes de couleur sur un même graphe visibles en même temps et indiquant la densité optique de la lumière émise par le gel à chaque longueur d’onde spécifique d’un des quatre ddN*, en fonction du temps de migration qui est proportionnel à la longueur des fragments qui migrent. Les fragments successifs apparaissent comme autant de pics qui se succèdent, le suivant ayant toujours un nucléotide de plus que le précédent. L’ordinateur interprète alors chacun de ces pics en fonction de la longueur d’onde et donne le résultat sous forme d’une séquence qui s’écrit au dessus du graphe. Dans une autre fenètre de l’écran la séquence obtenue peut-être alignée avec les autres séquences du même gène lues sur le même gel ou déjà enregistrées comme séquences de référence.
- Sur cet exemple, le pointeur de la souris désigne un moment de l’électrophorèse où sont passés successivement deux pics représentés en bleu, donc deux C sur la séquence. Le séquençage a été fait sur les deux brins de l’ADN et les séquences obtenues sont identiques. L’alignement de ces deux séquences avec une séquence de référence fait apparaître (position 410) que cette dernière est TCTCTT, alors que les séquences lues sur le gel sont TCTCCT. Il y a donc une substitution T$\rightarrow$C.



9.23 Séquençage : analyse de l'image

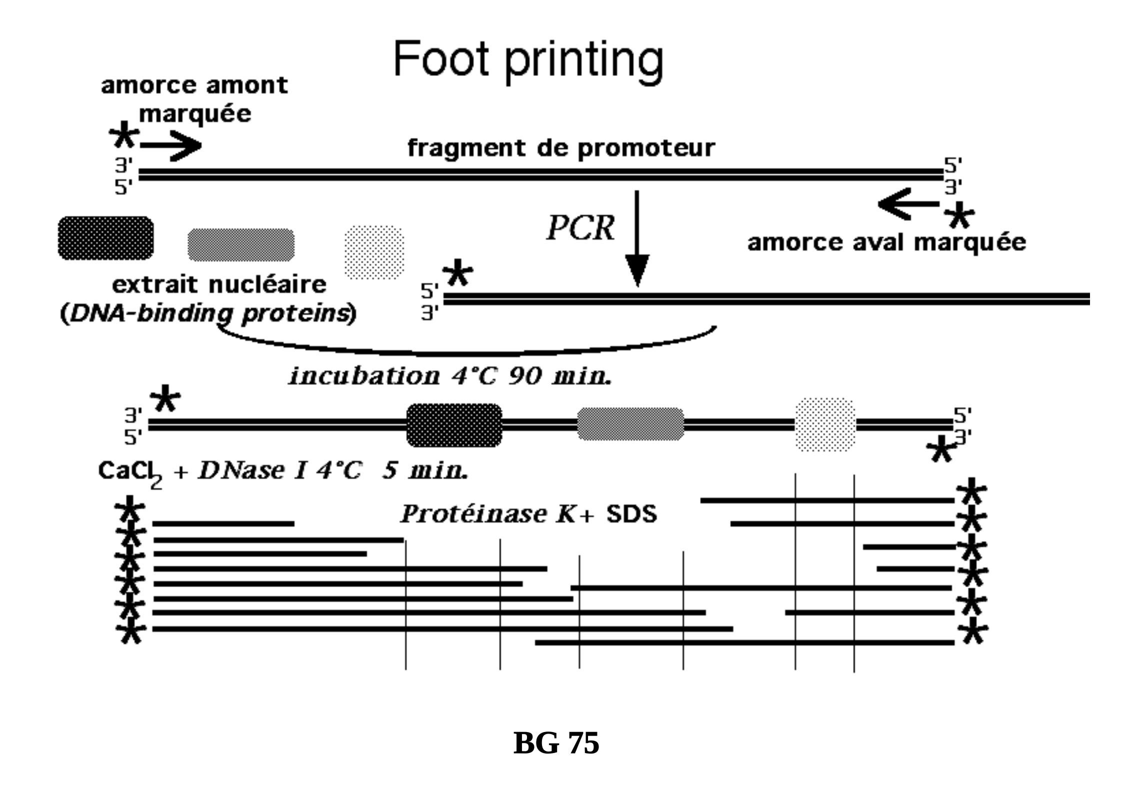


BG 74/2

- On rencontre parfois dans un séquençage deux pics superposés de couleurs différentes, ce qui crée une ambiguïté que l'ordinateur traduit par un N sur la séquence. Dans cet exemple, on voit à la même position un pic bleu et un pic rouge, donc un C ou un T. L'ADN génomique qui a été déposé provient d'un gène hétérozygote et chacun des deux gènes du même locus porte l'un un C et l'autre un T, de sorte qu'après amplification on obtient 50 % d'amplicons avec chacun de ces nucléotides.
- Pour distinguer les deux gènes de chaque locus hétérozygote, il faut auparavant faire un clonage spécifique qui permet de n'amplifier qu'un seul des deux amplicons de ces deux gènes.



9.24 Promoteur : *foot printing*

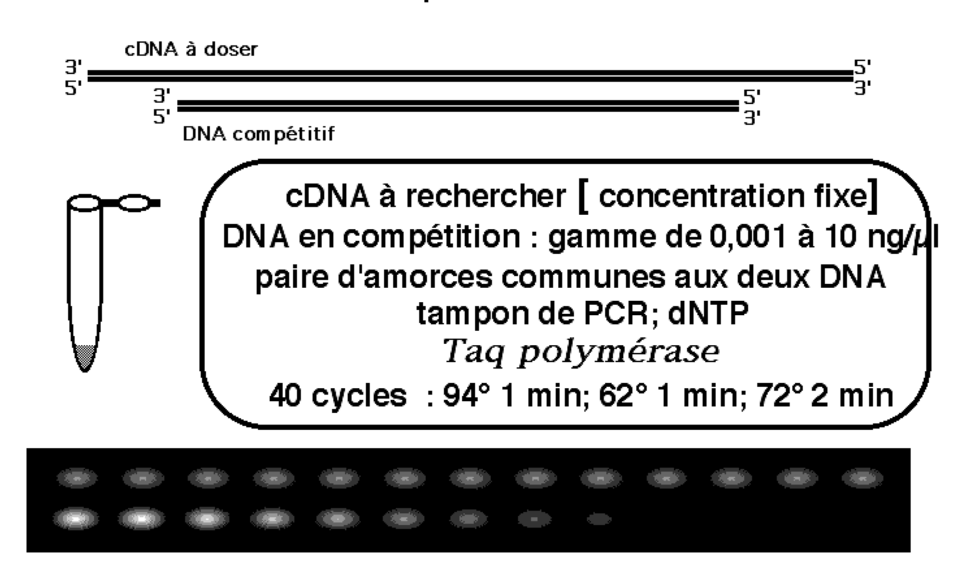


BG 75

- Dans l'analyse de la séquence des promoteurs, il est important de pouvoir caractériser les séquences des éléments cis-régulateurs.
- Des fragments de l'ADN du promoteur sont amplifiés par PCR et marqués. Ces fragments sont incubés en présence d'un extrait de noyaux (nucléoplasme) des cellules qui expriment le transcrit situé en aval de ce promoteur. Cet extrait contient les protéines capables de se lier à l'ADN (*DNA binding proteins*) parmi lesquelles se trouvent les éléments trans-régulateurs. Au cours de cette incubation la liaison entre les éléments trans- et cis- s'établit.
- L'ADN ainsi protégé par les éléments trans- est soumis à une digestion douce par la DNase I en présence d'ions Ca^{++}. La réaction est stoppée avec de l'EDTA et l'ADN ainsi digéré est réextrait (déprotéinisé) et la longueur des fragments est analysée par un séquenceur.
- Les fragments amplifiés ont été digérés au hasard par la DNase mais aucune coupure n'a pu se faire dans les séquences protégées de la DNase par les protéines nucléaires. Sur l'analyse les fragments se terminant dans ces zones ayant fixé les protéines, n'apparaîtront donc pas. Sur l'image du gel des plages noires (traces de pas = *foot prints*) marqueront les parties de la séquence qui ont fixé des protéines et sont donc probablement cis-régulatrices.



9.25 Dosage d'ARN : PCR quantitative

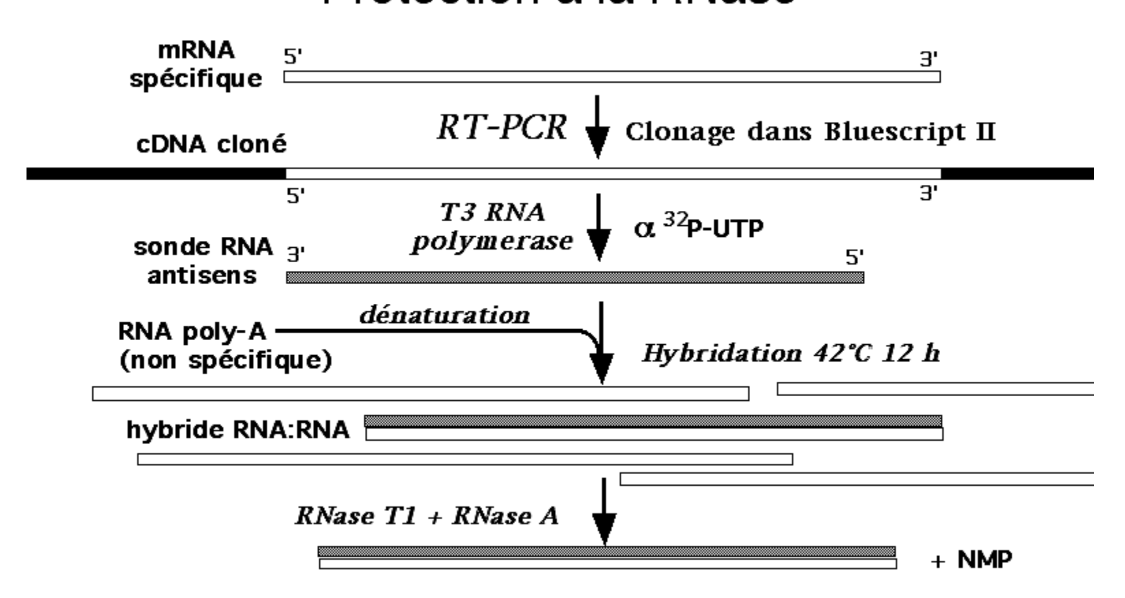


BG 76

- Pour mesurer l'expression d'un gène dans une cellule, on doit doser le taux du messager correspondant dans le cytoplasme. Ce taux dépend de la régulation de la transcription, mais aussi de la durée de vie du messager.
- Les différents mRNA sont amplifiés par une RT-PCR qui donne des cDNA dont les taux relatifs sont équivalents à ceux des mRNA qui ont servi de modèles.
- Afin de mesurer spécifiquement un messager précis on construit un DNA identique à ce messager, reconnaissable aux extrémités par les mêmes amorces, mais présentant une délétion au centre de la séquence le rendant plus court que le cDNA à rechercher. Dans une série de PCR on introduit le cDNA à doser, le DNA compétitif à des taux croissants pour faire une gamme de dosage, les amorces communes aux deux DNA et les autres facteurs pour la réaction.
- Dans l'électrophorèse qui suit, les cDNA amplifiés de concentration inconnue migrent moins que les DNA compétitifs de concentrations connues pour chaque puits. Les concentrations restant proportionnelles au cours de la PCR, la concentration du cDNA à doser est celle qui correspond à la concentration du DNA compétitif dans le tube où les deux taches de BET sont d'intensités égales.
- Pour pouvoir comparer les dosages des mRNA on fait en même temps (comme étalon interne) le dosage du mRNA d'un gène d'expression ubiquitaire et d'expression constante, par exemple actine ou facteur TFIID.



9.26 Dosage d’ARN : protection à la RNase



BG 77

- Pour détecter et doser un mRNA connu parmi les RNA-poly(A) d’un tissu, on peut faire une hybridation spécifique avec une sonde marquée.
- La sonde est un RNA complémentaire du messager à rechercher. Elle peut être construite à partir d’un mRNA connu dont le cDNA, après RT-PCR, a été cloné dans un phage. L’insert « sens » sert de modèle à une T3 RNA polymérase qui va synthétiser un RNA « antisens » complémentaire du mRNA de départ. Cette sonde « antisens » sera marquée (radioactive).
- 100 µg de RNA-poly(A) du tissu où la séquence est à rechercher sont incubés avec cette sonde dans des conditions d’hybridation lente en milieu liquide. Il se forme des hybrides RNA:RNA entre le mRNA recherché et la sonde antisens.
- Le milieu hybridé est alors incubé avec un mélange de ribonucléases toutes spécifiques du RNA simple brin, de telle sorte que seuls les hybrides formés avec la sonde peuvent résister à cette digestion.
- Le produit de digestion est ensuite déprotéinisé et repurifié, avant d’être analysé par électrophorèse. Les RNA hybrides marqués sont détectés par autoradiographie.




faculté de médecine
Pierre et Marie CURIE

Partie III

Caractérisation des événements génétiques

- Décrire et interpréter les résultats d'analyse (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) d'une lésion moléculaire conduisant à une expression anormale du gène considéré : mutation faux-sens, non-sens, décalage du cadre de lecture, protéine tronquée, …
- Expliquer les étapes (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) des techniques ayant permis : un diagnostic génotypique, une analyse génique ou une empreinte génétique : RFLP, carte de restriction, haplotypes, séquençage.




faculté de médecine
Pierre et Marie CURIE

Chapitre 10

Mutations



10.1 Polymorphismes de restriction

Polymorphismes de restriction

amorce 5' intron 17 exon 18

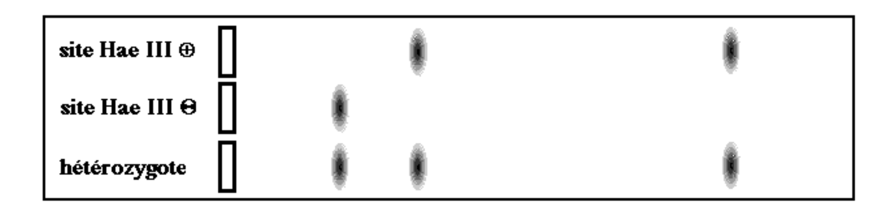
5' tccgctgtttaccatttgttggcag agACAGATGGTCAGTCTGGAGGAT

GACGTGGCG TGAacatctgcctggagtcccgcccctgcccagaaccctt

cctgagacctcgcc ggccttgtttattcaaacacagagaagaccaaagc3'

Hae III amorce 3'

tccgctgtttaccatttgttggcag agACAGATGGTCAGTCTGGAGGAT
GACGTGGCG TGAacatctgcctggagtcccgcccctgcccagaaccctt
cctgagacctcgcc gg ccttgtttattcaaacacagagaagaccaaagc

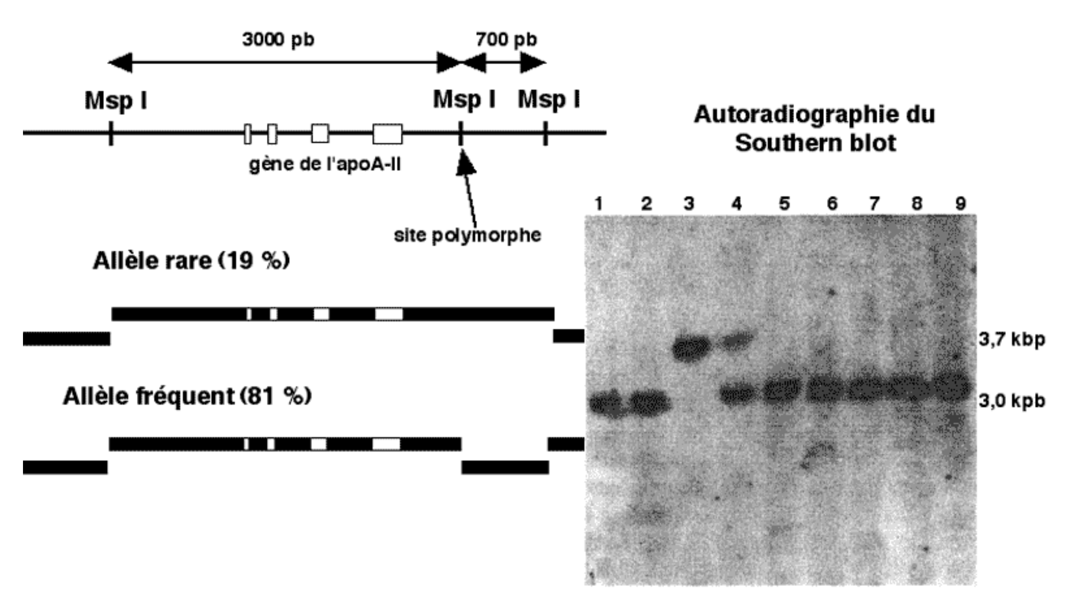


BG 78

- Les polymorphismes de longueur des fragments de restriction (*Restriction Length Fragment Polymorphism* = RFLP) sont des techniques simples pour analyser la séquence des acides nucléiques.
- Dans une séquence amplifiée par PCR à partir du gène du récepteur des LDL, comme celle (148 pb) représentée ci-dessus, on trouve des sites de restriction : par exemple Hae III (gg/cc).
- En digérant les amplicons par cette enzyme on peut distinguer les ADN ayant ce site (à l'électrophorèse deux fragments : 35 et 113 pb) de ceux ayant une séquence différente au niveau de ces quatre nucléotides (pas de coupure, un seul fragment de 148 pb). Si le prélèvement provient d'ADN génomique, les séquences amplifiées à partir des deux chromosomes homologues peuvent être différentes (hétérozygotie avec trois fragments : 35, 113 et 148 pb).
- Les polymorphismes sont fréquents dans la population : pour celui-ci il y a 77 % qui ont ce site Hae III (séquence « sauvage ») et 23 % qui ne l'ont pas.
- On peut aussi rechercher par cette technique des mutations, par définition beaucoup plus rares que ces polymorphismes.



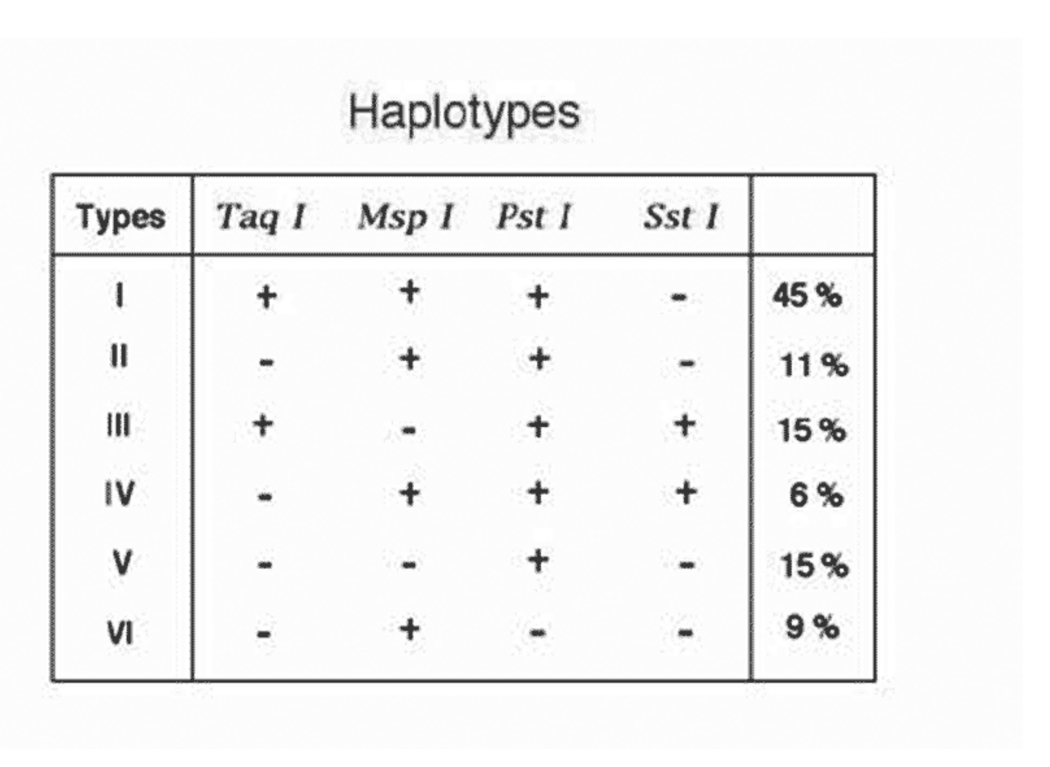
10.2 Polymorphisme Msp I de l'apoA-II



BG 78/1

- Le Southern blot permet la détection des polymorphismes de longueur des fragments de restriction (RFLP ou restriction fragments length polymorphism).
- Il existe un polymorphisme dans le gène de l'apolipoprotéine A-II. Ce polymorphisme altère la séquence du site reconnu par l'enzyme de restriction Msp I situé en aval du gène, de telle sorte que 19 % des sujets (anglais) ne présente pas de site de restriction à cet endroit.
- Lorsqu'on digère l'ADN génomique de plusieurs individus choisis au hasard on obtient des fragments de taille différentes entre chaque point de coupure par l'enzyme. Après électrophorèse pour séparer ces fragments en fonction de leur taille et transfert sur une membrane, on hybride les fragments sur la membrane avec une sonde complémentaire des exons de l'apoA-II puis on fait l'autoradiograhie de la membrane pour révéler ces fragments.
- La plupart de ces sujets (81 %) qui possèdent trois sites de coupure autour du gène donnent un fragment de 3000 paires de nucléotides (3,0 kb).
- D'autres sujets, plus rares (19 %) qui ne possèdent pas le site de restriction ont un fragment plus grand = 3700 paires de nucléotides (3,7 kb). Le sujet dont l'ADN a été déposé dans l'électrophorèse au n°3 est dans ce cas.
- Il arrive qu'un sujet porte un allèle du type fréquent sur un chromosome et l'autre allèle sur l'autre chromosome. Il est alors hétérozygote pour le polymorphisme et à l'électrophorèse (n°4) on voit les deux fragments révélés par la sonde.

10.3 Haplotypes



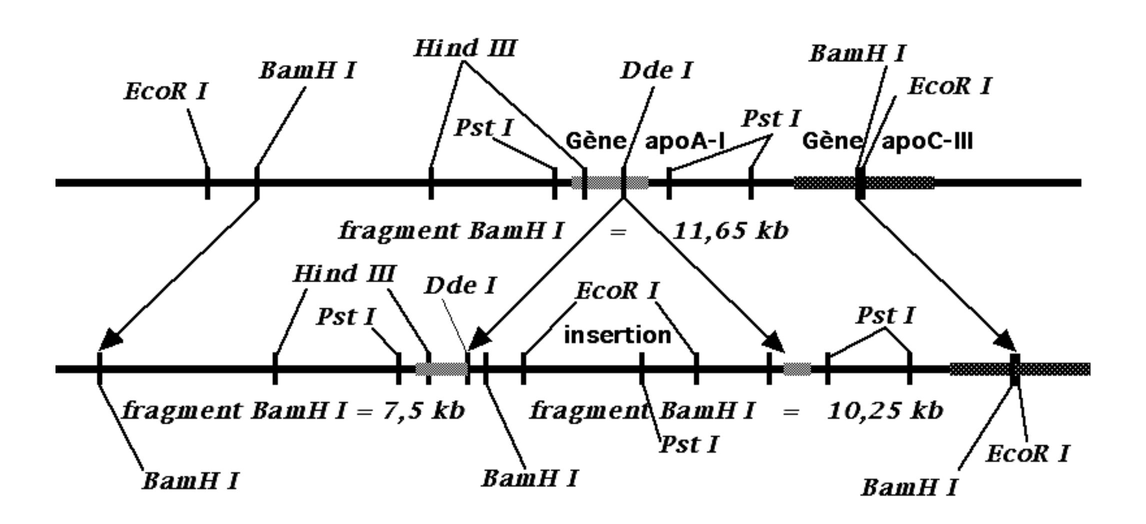
Haplotypes

Types	Taq I	Msp I	Pst I	Sst I	
I	+	+	+	-	45 %
II	-	+	+	-	11 %
III	+	-	+	+	15 %
IV	-	+	+	+	6 %
V	-	-	+	-	15 %
VI	-	+	-	-	9 %

BG 78/2

- Lorsqu'une maladie se transmet héréditairement dans une famille et que le gène responsable de cette maladie est unique et parfaitement identifié, les symptômes révèlent la présence d'au moins un allèle muté chez les personnes atteintes.
- Dans un grand nombre de familles le lien direct des symptômes avec une mutation connue du gène en cause n'est pas clairement établi. Dans ces cas là, on cherche à identifier l'allèle lié à la maladie par des marqueurs génétiques qui caractérisent les différents allèles du gène : polymorphismes, microsatellites de longueur variable, ... Un seul de ces marqueurs n'est pas suffisant pour repérer exactement chaque allèle, puisque plusieurs allèles peuvent avoir un marqueur en commun. Mais en associant plusieurs marqueurs on arrive à caractériser de façon précise chacun des allèles transmis dans la famille jusqu'à ce qu'un de ces allèles apparaisse comme exactement lié à l'apparition de la maladie. Les allèles ainsi caractérisés par plusieurs marqueurs génétiques sont appelés des haplotypes.
- Sur ce tableau sont indiqués les principaux haplotypes du groupe de gènes des apolipoprotéines du chromosome 11, caractérisés par quatre polymorphismes de restriction (Taq I, Msp I, Pst I et Sst I) situés de 5' en 3' dans un des deux brins de ce groupe de gènes. La présence (+) où l'absence (-) de chaque site de restriction est indiquée : ainsi un allèle caractérisé par la présence des 3 premiers sites et l'absence du site Sst I est présent chez 45 % des sujets de cette population, alors que celui qui ne possède que le seul site Msp I n'est retrouvé que chez 9 % des sujets.

10.4 Cartes de restriction



BG 79

- Les marqueurs génétiques peuvent aussi caractériser les ADN responsables des maladies héréditaires grâce aux cartes de restriction.
- L'ADN du malade est digéré par de nombreuses enzymes de restriction isolément ou associées entre elles. Les fragments de restriction sont reconnus par une sonde spécifique du gène responsable de la maladie. Il en résulte un grand nombre de fragments possibles dont les longueurs (en kilobases) sont mesurées par électrophorèse à coté d'un marqueur de masse moléculaire.
- Ainsi, le DNA d'un malade atteint d'une dyslipoprotéinémie et celui d'un témoin sain ont été digérés par BamH I : le sujet sain montre un seul fragment long de 11,65 kb reconnaissable par la sonde du gène de l'apoA-I, alors que l'ADN du malade en montre deux de 7,5 et 10,25 kb. Le fragment BamH I du sujet témoin peut encore être digéré par d'autres enzymes donnant à chaque fois deux fragments (Dde I) ou trois (Hind III) ou quatre (Pst I).
- Les fragments peuvent être comparés comme les morceaux d'un puzzle pour obtenir une carte des emplacements des sites de restriction sur l'ADN. La même carte, chez le sujet malade, montre les mêmes sites de restriction aux deux extrémités du fragment BamH I, mais il apparaît une insertion de 6 kb au milieu du gène de l'apoA-I avec des sites nouveaux : BamH I, EcoR I, Pst I non retrouvés chez le témoin. Cette insertion est responsable d'une absence d'expression de l'apoA-I ce qui se traduit par une dyslipoprotéinémie.




faculté de médecine
Pierre et Marie CURIE

Chapitre 11

Expression



11.1 Vecteur (définition)

VECTEUR :

- **Structure biologique capable de complexer une macromolécule et de l'intégrer spécifiquement dans une cellule vivante.**

BG 80

- Un vecteur est un transporteur capable de conduire une molécule à pénétrer dans une cellule alors qu'elle n'est pas captée spontanément par cette cellule.
- Les vecteurs sont principalement étudiés pour transporter des médicaments dans un type déterminé de cellules de l'organisme d'un malade, tout en mettant ce médicament à l'abri des enzymes de détoxification du sujet.
- En biologie moléculaire, certains vecteurs sont étudiés pour faire entrer un acide nucléique dans une cellule et permettre la réplication ou l'expression de cet acide nucléique dans la cellule qui le reçoit.
- Les plasmides bactériens, les bactériophages et de nombreux virus sont des vecteurs naturels permettant la transfection de différents types de cellules-hôtes. D'autres vecteurs encore plus performants ont été obtenus artificiellement à partir de ceux-là grâce à des manipulations génétiques.



11.2 Cellules-hôtes

Cellules-hôtes

Escherichia coli

DH5α : expression de pUC19
HB 101 : expression de pBR322
JM109 : lac œ pour sélection

Saccharomyces cerevisiæ
Saccharomyces pompe
Pichia Pastoris

Spodoptera frugiperda
Trichoplusiani

Chinese Hamster Ovaries (CHO)
Baby hamster kidney (BHK)
cellules HeLa
lignée 3T3

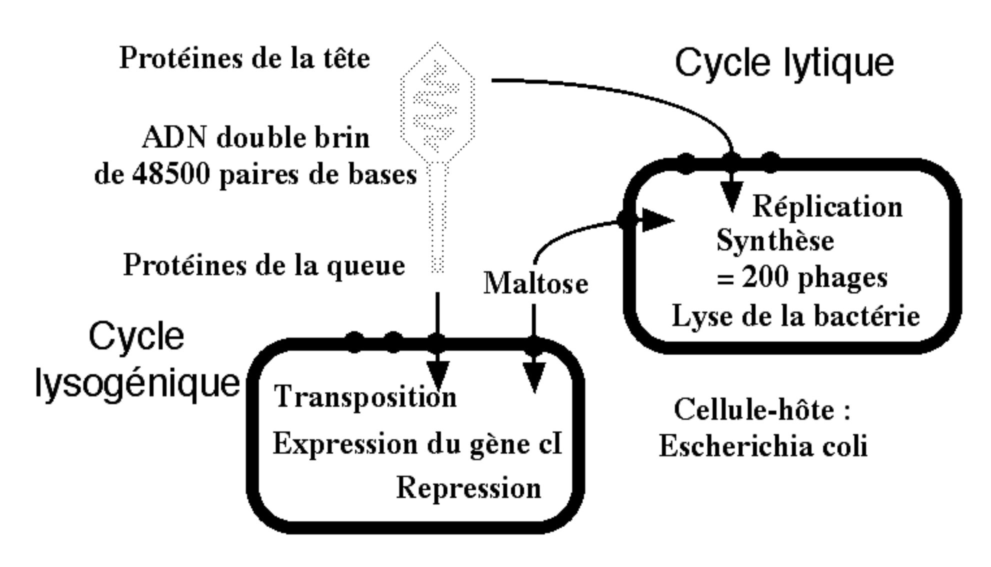
Récepteur de l'ecdysone (Drosophile)
+ Promoteur activé par l'ecdysone

BG 81

- Les cellules qui sont habituellement transfectées par des vecteurs contenant des fragments d'ADN, sont destinées à permettre d'amplifier ce vecteur en même temps que leur croissance.
- Des bactéries comme *Escherichia coli* sont utilisées pour recevoir des plasmides. Certaines souches de cette espèce sont plus particulièrement développées pour permettre la transfection de certain vecteurs : souche DH5α pour les plasmides pUC18 ou pUC19, HB 101 pour le plasmide pBR322, souche JM 109 dépourvue d'opéron lactose pour les vecteurs à β-galactosidase.
- Des levures comme *Saccharomyces cerevisiae* (levure de bière) ou *Pichia pastoris* sont utilisées pour des vecteurs eucaryotiques, ainsi que des végétaux : *Spodoptera frugiperda* ou *Trichoplusiani*.
- Les cellules animales en culture du hamster doré : CHO (*Chinese hamster ovaries*) ou BHK (*baby hamster kidneys*) ou de la souris : lignée 3T3 peuvent recevoir des vecteurs spécifiques. Certaines lignées de cellules humaines en culture (cellules HeLa) peuvent aussi être employées.
- Certains vecteurs sont employés pour exprimer des gènes commandés par des promoteurs spécifiques : promoteur des métallothionéïnes, répondant à la présence de métaux lourds (Cd par exemple) ou promoteur induit par l'ecdysone (hormone stéroïde des Insectes).



11.3 Cycle du phage λ

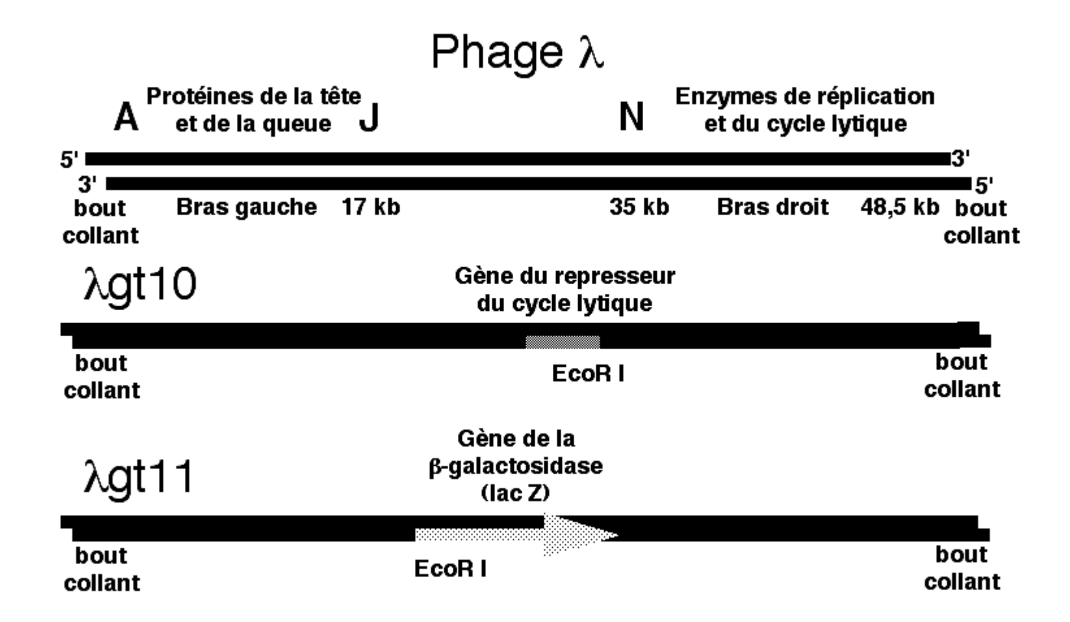


BG 82

- Le génome du bactériophage λ dans les particules virales est représenté par un ADN double-brin de 48,5 kb, dont les extrémités sont cohésives entre elles (bouts collants). Le virus entre dans les cellules bactériennes par un récepteur (lamB) qui sert normalement au transport du maltose. L'expression de ce récepteur est induite par le maltose et réprimée par le glucose.
- Dès qu'il pénètre dans la bactérie, l'ADN du phage se circularise par hybridation de ses extrémités et la DNA ligase de la bactérie achève de fermer les brèches. L'évolution se poursuit dans deux directions différentes en fonction des promoteurs activés de l'ADN du phage.
- La plupart du temps, l'expression des premiers gènes déclenche la réplication intense de l'ADN du phage et l'expression de protéines permettant la synthèse de particules virales. La bactérie meurt, éclate et libère environ 200 nouvelles particules virales (cycle lytique).
- Beaucoup plus rarement, le promoteur activé au départ transcrit le gène cI, qui réprime le cycle lytique et commande la transposition de l'ADN du phage dans le chromosome de la bactérie où il reste peu exprimé (prophage), mais répliqué à chaque division de la cellule-hôte (cycle lysogénique). Une élévation de la température déclenchera à nouveau l'excision du prophage et sa réplication comme dans le cycle lytique.



11.4 Phage λ

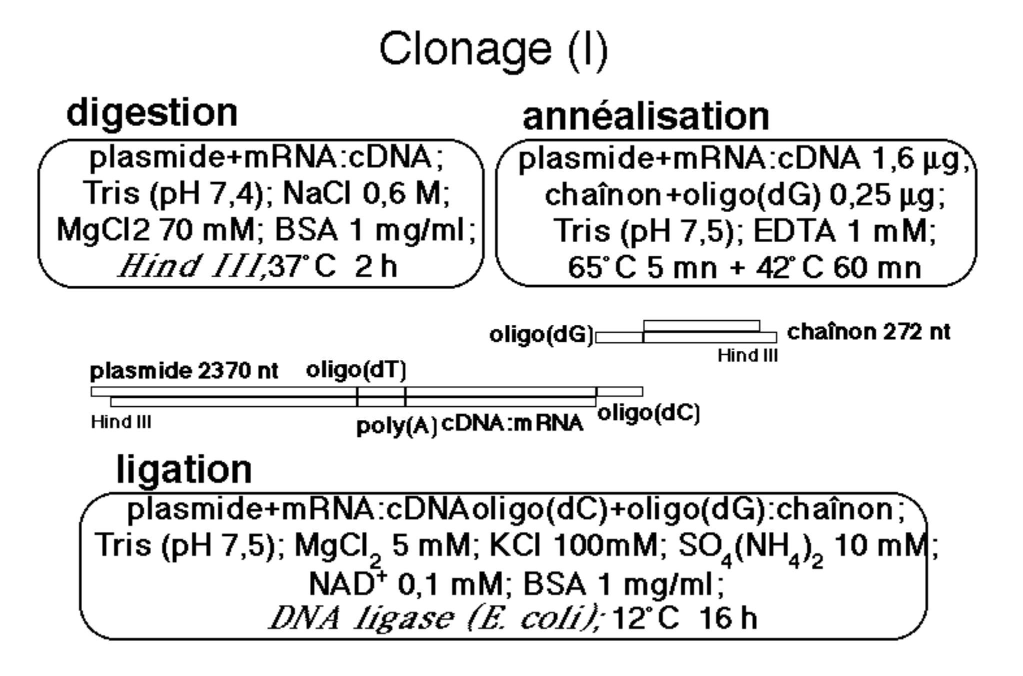


BG 83

- L’ADN du phage λ est un ADN linéaire de 48502 paires de nucléotides. Aux deux extrémités les brins sont inégaux mais leurs séquences sont complémentaires (bouts collants) de telle sorte que l’ADN du phage peut devenir circulaire.
- Les gènes transcrits à partir de cet ADN sont désignés par des lettres de l’alphabet : de A à J du côté 5’ terminal et de N à R dans la partie 3’ terminale. Il existe plusieurs promoteurs activés dans le cycle lytique : R1 qui transcrit de 5’ vers 3’ les gènes O à Q, R2 qui transcrit de S à J lorsque l’ADN est circulaire et que les brèches ont été réparées, ou bien dans le cycle lysogénique les promoteurs L1 et L2 qui transcrivent sur l’autre brin les gènes de la partie centrale.
- Le génome du phage λ (λ gt10) peut être modifié en supprimant le gène du represseur du cycle lytique dans la partie centrale, ce qui favorise la réplication du phage, et en ajoutant le gène de la β-galactosidase après le gène J (λ gt11)



11.5 Clonage (I)

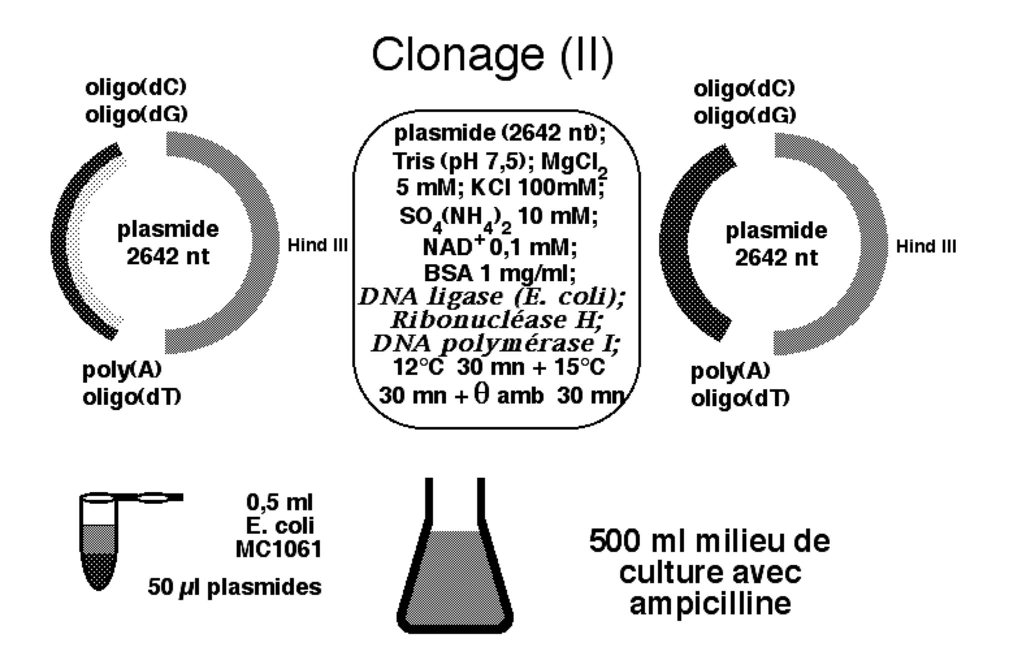


BG 84

- Un plasmide est utilisé pour insérer un ARN messager spécifique. Une extrémité à bouts francs a été prolongée par une queue poly(dT) grâce à la terminal transférase. Cet oligo(dT) en s'hybridant avec la queue poly(A) du messager, fournit un site d'initiation à une DNA polymérase pour synthétiser un cDNA complémentaire du messager initial. A l'extrémité 3'OH de ce fragment, une nouvelle queue poly(dC) est encore ajoutée par une terminal transférase. Le montage est alors digéré par Hind III qui hydrolyse un site à l'autre extrémité du plasmide (digestion).
- On prépare aussi un chaînon de ligation pourvu d'une queue poly(dG) par la terminal transférase et digéré de l'autre côté par Hind III. En incubant ensemble ces deux fragments d'ADN on forme un ADN circulaire (annéalisation) par hybridation des deux extrémités Hind III et des queues oligo(dC) et oligo (dG).
- L'ADN circulaire ainsi obtenu sera enfin incubé avec la DNA ligase pour fermer les deux brèches du site Hind III et la liaison entre l'extrémité 3' de la queue poly(dC) et le chaînon intermédiaire (ligation).



11.6 Clonage (II)



BG 84/1

- On est alors en présence d'un ADN circulaire comportant à droite la séquence du plasmide dont le site Hind III a été refermé et une insertion comprenant un mRNA (en vert) entre la fin de l'oligo(dG) et la fin de la queue poly(A) et le DNA complémentaire (cDNA, en jaune) entre le début de l'oligo(dT) et la fin de l'oligo(dC).
- L'ARN inséré est alors digéré par la RNase H, spécifique des hybrides RNA:DNA. Cette digestion crée une brèche qui sera comblée par une DNA polymérase et enfin fermée par la DNA ligase.
- On obtient alors un plasmide fonctionnel d'ADN circulaire dans lequel est inséré un ADN double brin correspondant à la séquence du messager. Ce plasmide est utilisé pour transfecter une souche compétente de bactéries-hôtes auxquelles il confère un gène de résistance à l'ampicilline.
- La culture de cette bactérie, en présence de l'antibiotique, permet de récupérer une grande quantité de cet ADN circulaire qui peut être utilisé ensuite pour caractériser, séquencer ou exprimer la séquence clonée du messager de départ.



11.7 EMBL

EMBL

EMBL 3

Sal I BamH I EcoR I

34 kb

bout collant

20 kb

EcoR I BamH I Sal I

bout collant

EMBL 4

EcoR I BamH I Sal I

34 kb

bout collant

20 kb

Sal I BamH I EcoR I

bout collant

BG 85

- Les phages EMBL3 et EMBL4 sont aussi des phages λ modifiés. Leur ADN fait 42364 paires de bases.
- Des emplacements spécifiques ont été préparés dans ces phages pour pouvoir y introduire des fragments d’ADN à amplifier. Pour EMBL3, un de ces emplacements se situe après le gène J et comprend les uns à la suite des autres trois sites de restriction : Sal I, BamH I et EcoR I ; l’autre emplacement se situe sur le brin antiparallèle et comprend les mêmes sites de restriction.
- Pour EMBL4 l’organisation est la même mais les sites de restriction sont dans l’ordre inverse : EcoR I, BamH I et Sal I.
- L’intérêt de tels sites est de permettre l’incorporation d’un ADN en contrôlant le sens de la transposition. Imaginez un fragment d’ADN limité d’un coté par Sal I et de l’autre par EcoR I. En digérant le phage par les mêmes enzymes puis en recombinant le phage avec ce fragment d’ADN on obtiendra une transposition dans un sens pour EMBL3 et dans l’autre sens pour EMBL4 (clonage directionnel).

11.8 *Polylinker*

Polylinker

```
M13mp18

        Sst I     Sma I
       = Sac I    Xma I     Xba I       Pst I     Hind III
ACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGCA
TGCTTAAGCTCGAGCCATGGGCCCCTAGGAGATCTCAGCTGGACGTCCGTACGTTCGAACCGT
  EcoR I     Kpn I     BamH I     Sal I       Sph I
                                  Acc I
                                 Hinc II
                                = Hind II

M13mp19
                                         Sma I     Sst I
    Hind III      Pst I      Xba I       Xma I    = Sac I
ACGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCA
TGCGGTTCGAACGTACGGACGTCCAGCTGAGATCTCCTAGGGGCCCATGGCTCGAGCTTAAGT
            Sph I      Sal I      BamH I   Kpn I      EcoR I
                       Acc I
                      Hinc II
                     = Hind II
```

BG 86

- L’avantage de ce clonage directionnel ne peut être obtenu qu’avec des sites de restriction uniques et bien placés dans le fragment d’ADN qu’on veut transposer dans le vecteur.
- Afin de disposer d’un maximum de possibilités, on construit des vecteurs dits de clonage, qui comportent des domaines riches en sites uniques de restriction et placés dans l’ordre inverse pour chaque paire de vecteurs. Ces domaines permettant des liaisons sur de multiples sites sont appelés « *polylinkers* ».
- Ainsi la paire de plasmides M13mp18 et M13mp19, comporte dans leur structure des domaines formés de l’association successive de dix sites de restriction dans un sens et dans l’autre.



11.9 Clonage directionnel

Clonage directionnel

...GTACCCGGG GATCCTCTAGAG TCGACCTGC... Sal I
...CATGGGCCCCTAG GAGATCTCAGCT GGACG... vecteur 1
BamH I
...CAGG TCGACTCTAGAG GATCCCCGG...
...GTCCAGCT GAGATCTCCTAG GGGCC... vecteur 2

ADN à insérer ...G GATCCNNN...NNG TCGAC...
...CCTAG GNNN...NNCAGCT G...

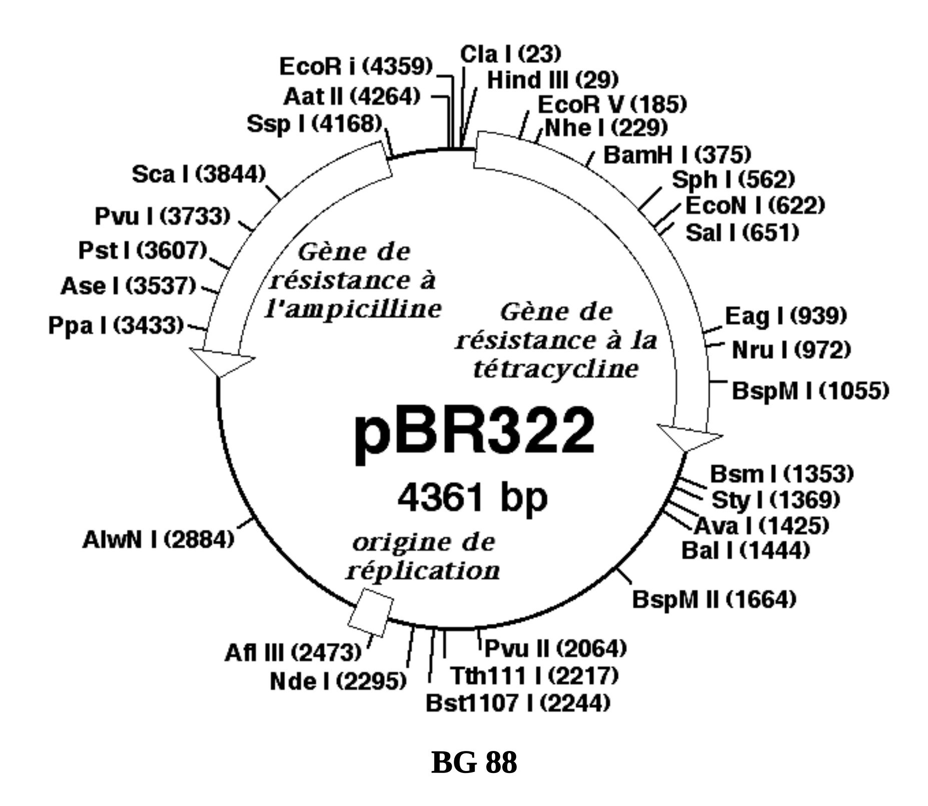
...GTACCCGGGGATCCNNN...NNGTCGACCTGC... ADN inséré
...CATGGGCCCCTAGGNNN...NNCAGCTGGACG... dans vecteur 1
à l'endroit
et
...CAGGTCGACNN...NNNGGATCCCCGG... dans vecteur 2
...GTCCAGCTGNN...NNNCCTAGGGGCC... à l'envers

BG 87

- L'insertion d'un fragment d'ADN dans un *polylinker* permet d'orienter le fragment par rapport au vecteur et d'obtenir par exemple le séquençage d'un fragement dans chaque sens.
- Supposons un fragment d'ADN double brin digéré successivement par les enzymes Sal I et BamH I de telle façon que les extrémités soient cohésives (bouts collants de chacun de ces sites de restriction). On utilise pour faire le clonage une paire de vecteurs pourvu d'un *polylinker* dans un sens pour le vecteur 1 et en sens inverse pour le vecteur 2. Lors de la recombinaison le fragment d'ADN vient s'insérer dans un sens sur le vecteur 1 et dans l'autre sur le vecteur 2. Après la croissance bactérienne on dispose de colonies n'ayant un insert que dans un seul sens. Grâce à des amorces spécifiques de l'ADN du vecteur on peut alors séquencer le brin d'ADN spécifiquement dans le sens où il est inséré.
- Cette technique est utilisée principalement pour effectuer le séquençage des gènes sur les deux brins, la séquence antiparallèle et complémentaire permettant de contrôler la séquence d'un premier brin.



11.10 Carte du plasmide pBR322

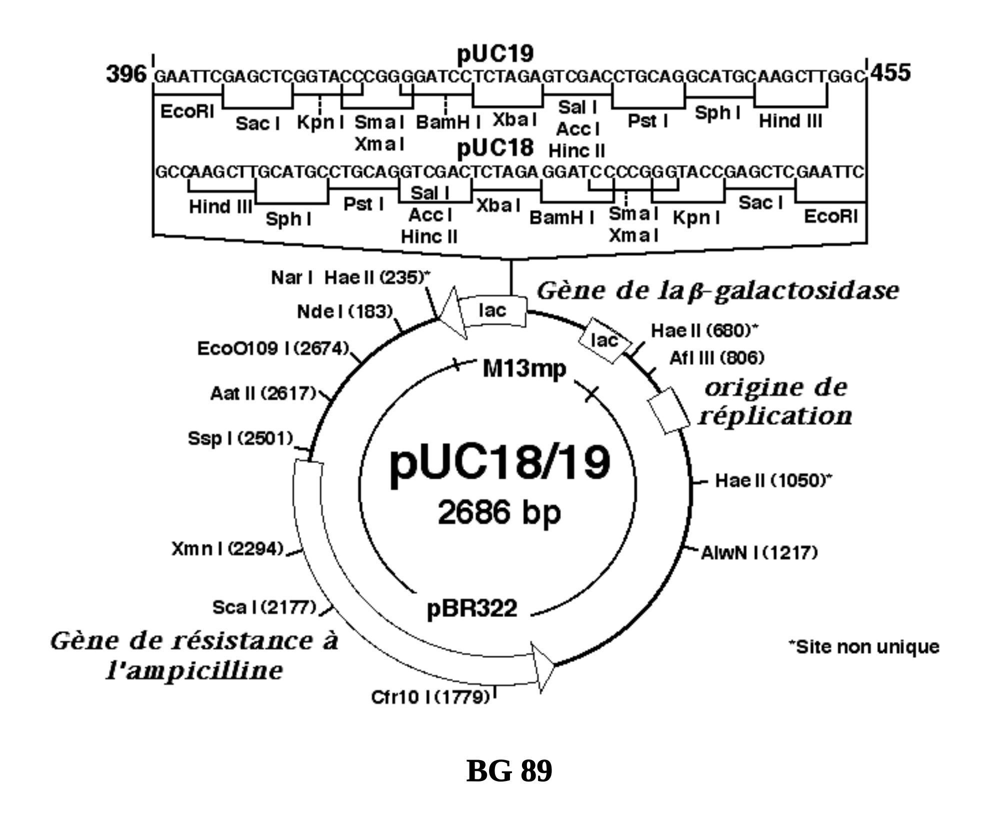


BG 88

- Le plasmide pBR322 est un ADN circulaire double brin de 4361 paires de bases. Par convention le nucléotide n°1 est situé au milieu du site de restriction EcoR I en direction du gène tet^r.
- Il contient un origine de réplication (2535), un gène de résistance à la tétracycline (tet^r 86-1276) et un gène de résistance à l'ampicilline (amp^r 4153-3293). Le gène amp^r code pour une protéine (β-lactamase) de 286 acides aminés capable de cataboliser cet antibiotique.
- Le plasmide contient de nombreux sites de restriction répartis sur toute la séquence. Beaucoup de ces sites sont uniques, permettant de transformer l'ADN circulaire en ADN linéaire. Exemple, le site BamH I (position 375) au début du gène tet^r.
- La digestion par BamH I permet de recombiner le vecteur avec un fragment de digestion par BamH I d'un ADN à cloner. La ligase du bactériophage T4 permet de recirculariser l'ADN du plasmide. Les cellules transfectées avec un tel plasmide recombinant feront la réplication du plasmide au cours de leur croissance. Elles n'expriment plus le gène tet^r, mais restent résistantes à l'ampicilline ce qui permet de les sélectionner et d'isoler l'ADN du plasmide ainsi cloné.



11.11 Carte des plasmides pUC18/19

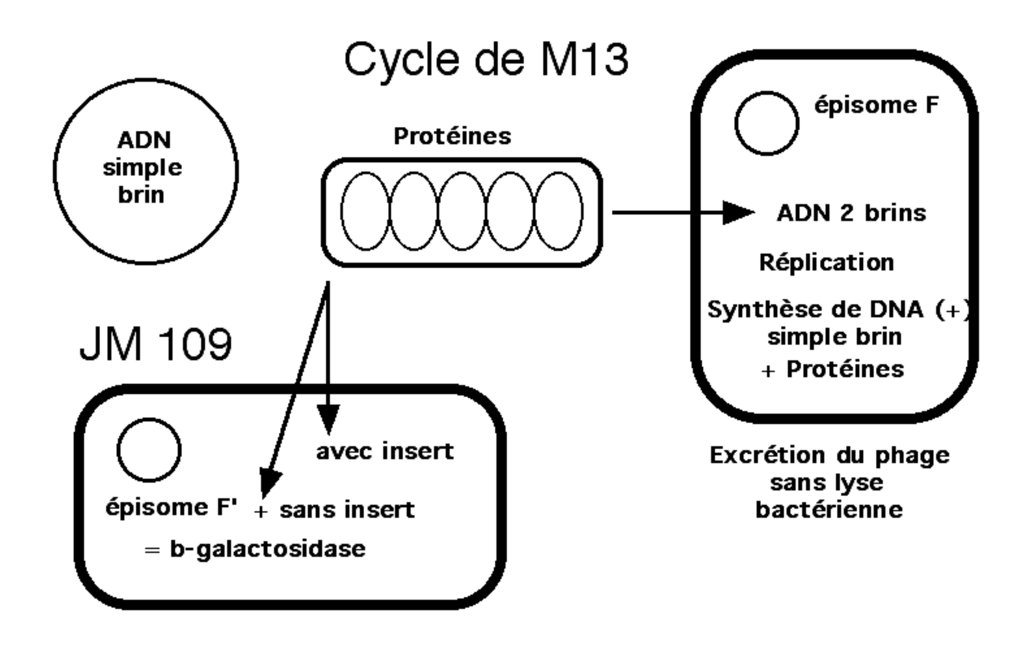


BG 89

- Les plasmides pUC18 et pUC19 sont des ADN circulaires double brin de 2696 paires de bases (pb). Ils ont été construits par recombinaison d'un fragment de pBR322 (2250 pb) et d'un fragment de M13 double brin (446 pb) : M13mp18 pour pUC18 et M13mp19 pour pUC19.
- L'ADN issu de pBR322 contient le gène amp^r qui permet la sélection par la résistance à l'ampicilline. L'ADN issu de M13 contient le promoteur du gène lacZ suivi d'une séquence modifiée comprenant un polylinker (voir BG 86, section 11.8 page 155) et la séquence de la partie NH2 terminale du gène de la β-galactosidase. Cette séquence est indispensable pour que la bactérie hôte puisse faire la réaction de cette enzyme (phénomène d'α-complémentation, voir BG 44 section 6.1 page 74).
- Le polylinker, hydrolysé par une des enzymes de restriction, peut recevoir une séquence d'ADN liée par recombinaison. Dans ce cas, l'expression du peptide complémentaire de la β-galactosidase est interrompue et la bactérie ne pourra plus digérer le X-Gal, ce qui l'empêchera de prendre la couleur bleue caractéristique. Les colonies résistantes à l'ampicilline et de couleur blanche sont donc transfectées par le vecteur recombinant.



11.12 Cycle de M13

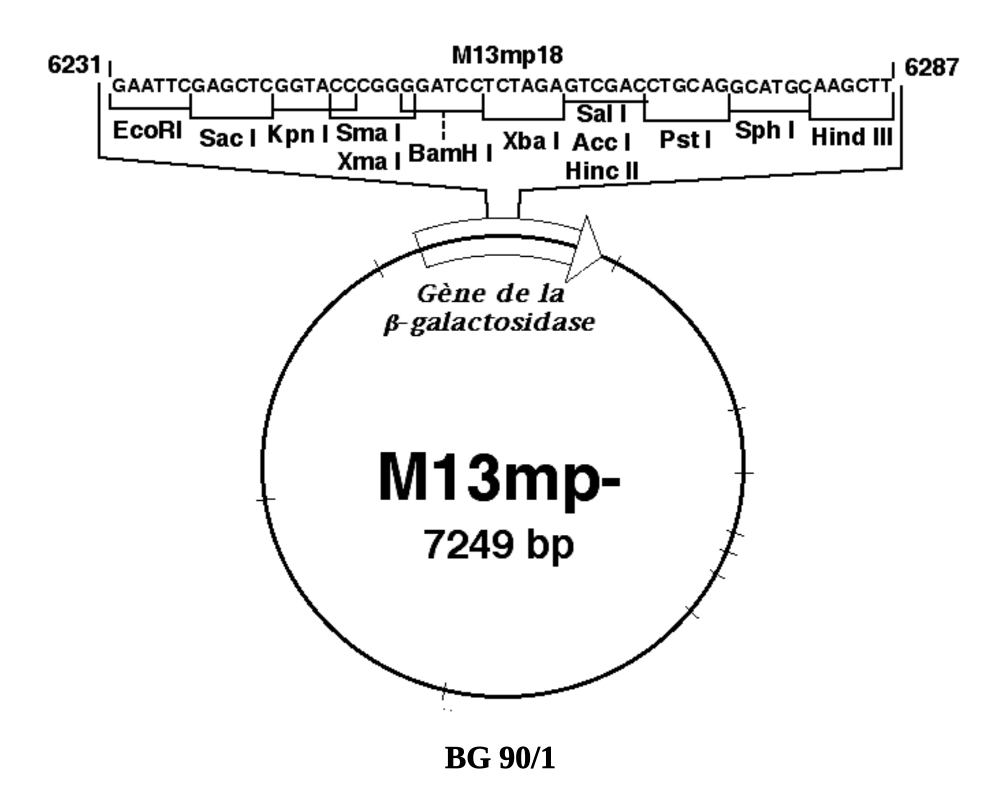


BG 90

- Le bactériophage M13 est un phage dont la forme libre est constituée d'une membrane entourant un ADN simple brin de 6,4 kilobases.
- La transfection des bactéries par M13 et l'expression du virus dépendent de la présence d'un chromosome surnuméraire bactérien (épisome F) : la présence de cet épisome permet le cycle normal du phage : synthèse d'un brin d'ADN complémentaire, réplication et en fin de cycle synthèse spécifique d'un seul des deux brins de l'ADN (brin +) qui sera pourvu d'une membrane et sécrété hors de la cellule.
- Lorsque l'épisome ne contient pas les gènes permettant ce cycle (épisome F'), l'ADN du phage s'incorpore sous la forme d'un ADN double brin capable d'exprimer les gènes qu'il contient ou qu'on a inséré dans sa structure, par exemple la β-galactosidase. Les bactéries avec l'épisome F' sont des variétés commercialisées pour permettre l'expression des gènes clonés dans le phage M13.
- Le phage M13 existant sous la forme d'un ADN simple brin, il est possible d'y insérer des brins d'ADN d'un gène particulier présent à l'état hétérozygote chez l'homme : dans ce cas chaque phage ne peut insérer qu'un seul des deux allèles et les deux allèles se retrouvent dans des colonies séparées.



11.13 Carte du M13



BG 90/1

- Le bactériophage M13 est constitué d'un ADN simple brin. On a construit des vecteurs en introduisant dans la séquence une forme tronquée du gène de la β-galactosidase d'*Escherichia coli* (gène LacZ'). L'ensemble a une longueur de 7249 paires de bases (pb).
- Le gène lacZ' code pour la partie NH_2 terminale du gène de la β-galactosidase. Cette séquence (peptide de complémentation) est indispensable pour que la bactérie hôte puisse faire la réaction de cette enzyme (phénomène d'α-complémentation, voir BG 44 section 6.1 page 74).
- Les codons 5 à 18 de ce gène lacZ' forment dans M13mp18 un ensemble de séquences spécifiques d'enzymes de restriction (sites de restrictions dans l'ADN double brin) qu'on nomme *polylinker* ou site de clonage multiple. Dans M13mp19 les mêmes sites de restriction se trouvent dans l'ordre inverse.
- Le *polylinker*, hydrolysé par une des enzymes de restriction, peut recevoir une séquence d'ADN liée par recombinaison. Dans ce cas, l'expression du peptide complémentaire de la β-galactosidase est interrompue et la bactérie ne pourra plus digérer le X-Gal, ce qui l'empêchera de prendre la couleur bleue caractéristique. Les colonies résistantes à l'ampicilline et de couleur blanche sont donc transfectées par le vecteur recombinant.



Partie IV

Le génie génétique

Rappel des objectifs

- Définir[1] les termes suivants : vecteur, clonage, polylinker, plasmide.
- Décrire les étapes de l'insertion d'un fragment d'ADN dans un plasmide.
- Expliquer les étapes (sur un exemple emprunté au cours, aux T.P. ou à votre expérience personnelle) des techniques ayant permis : une mutagénèse, la transposition d'un gène dans une bactérie, la création d'un animal ou d'une plante transgénique.
- Expliquer le principe de la construction d'un vecteur : *cet objectif servira à tester la faculté pour l'étudiant(e) de faire une synthèse des connaissances de l'ensemble des chapitres précedents.*

1. **Définir** : préciser dans une phrase concise l'essence d'un objet ou les limites d'un concept en excluant toute notion étrangère et en comprenant toutes les variations possibles de l'objet ou du concept cerné.




faculté de médecine
Pierre et Marie CURIE

Chapitre 12

Mutagénèse



12.1 Création d'un site de restriction

Normal

Brin sens

AAC ACT TAC TTG AAT TCC AAG AGC ACA CGG TCT TCA

Amorce

AAC ACT TAC TTG AAT TCC AAG AGC ACC C

TTG TGA ATG AAC TTA AGG TTC TCG TGT GCC AGA AGT

Brin antisens

Msp I

AAC ACT TAC TTG AAT TCC AAG AGC ACC CGG TCT TCA

Mutation

Brin sens

AAC ACT TAC TTG AAT TCC AAG AGC ACA CAG TCT TCA

Amorce

AAC ACT TAC TTG AAT TCC AAG AGC ACC C

TTG TGA ATG AAC TTA AGG TTC TCG TGT GTC AGA AGT

Brin antisens

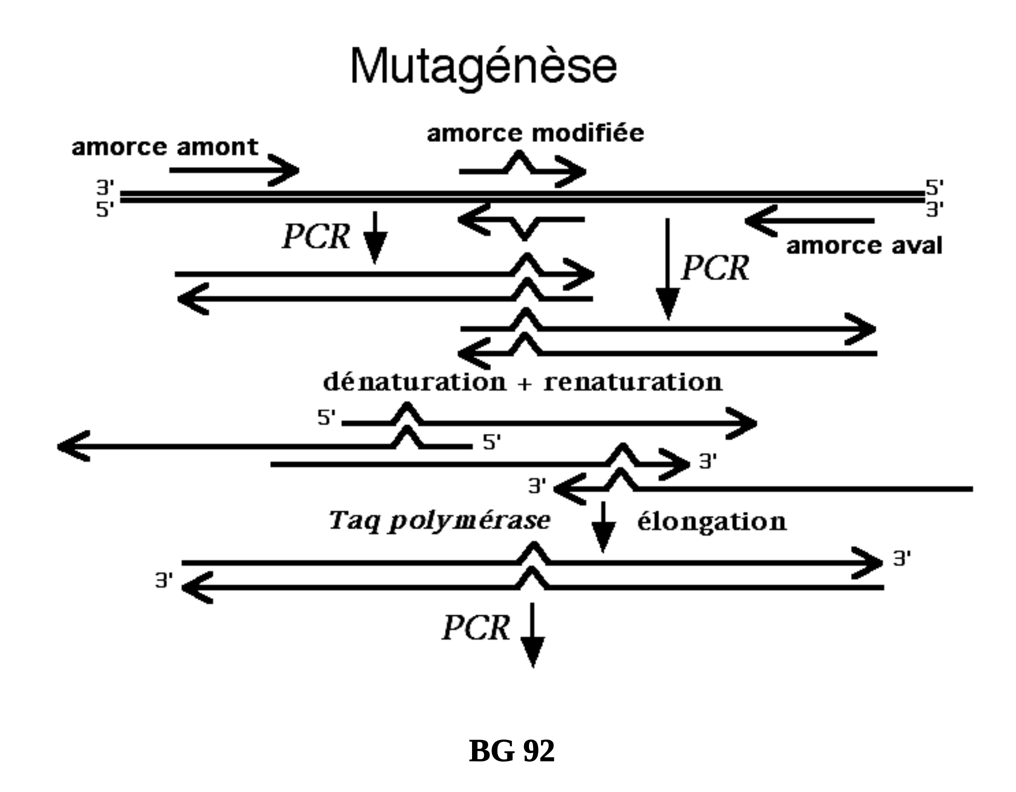
AAC ACT TAC TTG AAT TCC AAG AGC ACC CAG TCT TCA

BG 91

- Le polymorphisme des fragments de restriction est une des façons les plus simples de mettre en évidence une mutation dans l'ADN. Mais il arrive que la mutation ne modifie aucun des sites de restriction de l'ADN du gène. La mutation Arg3500→Gln de l'apolipoprotéine B en est un exemple.
- Pour permettre l'identification d'une telle mutation on va faire l'amplification grâce à des amorces modifiées pour créer un polymorphisme parmi les fragments de digestion de l'ADN amplifié.
- Pour amplifier le fragment de l'apoB on synthétise une amorce (côté 5') complémentaire du brin antisens sauf pour l'extrémité 3'OH qui est ACCC au lieu de ACAC. Le misappariement entre le C de l'amorce et le T du brin antisens n'est pas suffisant pour empêcher l'hybridation et donc la PCR se fait à partir de cette amorce : ACCCGGTCTTCA... Il apparaît donc dans l'ADN amplifié un site CCGG reconnu par l'enzyme Msp I qui digérera l'ADN à cet endroit.
- Si l'amplification se fait à partir de l'ADN muté : ACCCAGTCTTCA... le site Msp I n'apparaît pas, il n'y a donc pas d'hydrolyse à cet endroit et le fragment digéré par Msp I est plus long que dans l'ADN normal.



12.2 Mutagénèse par PCR



BG 92

- La création d'une mutation artificielle peut être obtenue dans un gène par le jeu des amplifications à partir d'amorces modifiées.
- On prépare l'amplification du gène ou du cDNA grâce à deux amorces placées en amont et en aval de la séquence d'intérêt. On prépare aussi par synthèse des oligonucléotides conformes aux séquences des deux brins du gène autour du codon qui doit être modifié mais dont les bases sont volontairement changées pour qu'elles permettent la synthèse d'un brin complémentaire comprenant la substitution (ou la délétion...) souhaitée.
- On fait une première amplification entre chacune des amorces modifiées et les amorces des extrémités. En réunissant ces deux amplifications et en dénaturant l'ADN, on conduit à une renaturation entre les fragments au niveau des amorces centrales modifiées. Cette hybridation fait apparaître des extrémités 3'OH d'où l'élongation (Taq polymérase) peut se poursuivre jusqu'aux extrémités de la séquence d'intérêt. La poursuite des cycles en présence des amorces amont et aval conduit à l'amplification de cette séquence complète dans laquelle le codon muté a été inséré.




faculté de médecine
Pierre et Marie CURIE

Chapitre 13

Transposition - recombinaison



13.1 Transposition (définition)

TRANSPOSITION

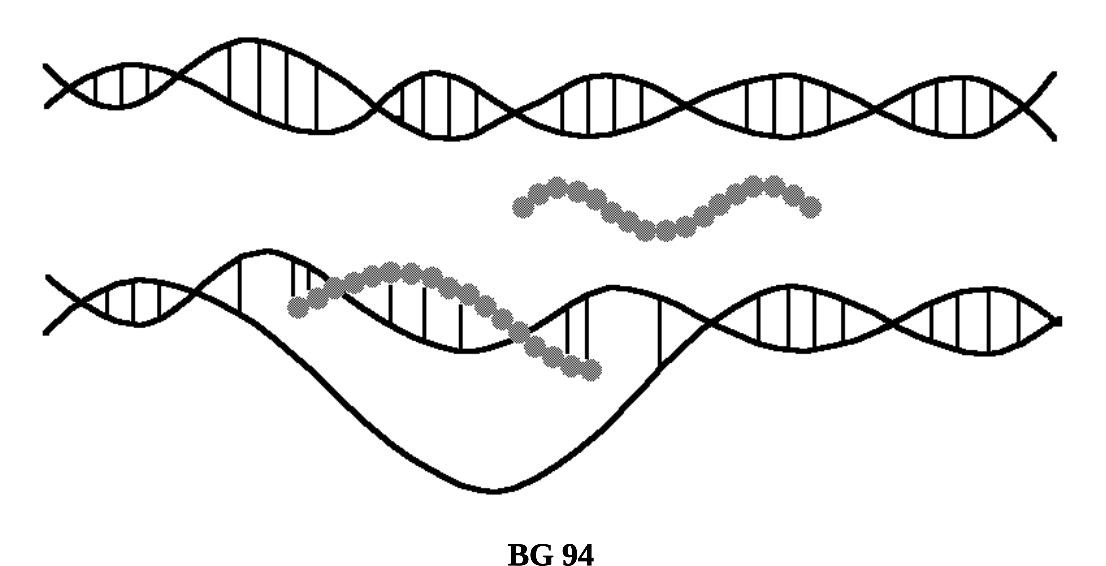
- **Incorporation d'une séquence d'acide désoxyribonucléique nouvelle dans une coupure du DNA génomique d'une cellule.**

BG 93

- Au cours de l'évolution des espèces, le génome s'agrandit continuellement par suite de duplications de gènes et de transpositions.
- Les duplications se produisent en tandem, une séquence se trouvant répétée deux fois dans le génome à la suite de la duplication. Cette duplication permet à chacune des copies de la séquence d'évoluer pour son propre compte ce qui augmente les chances de produire des caractères nouveaux.
- Les transpositions résultent de l'incorporation d'acides nucléiques synthétisés hors du génome :
  - soit par incorporation du DNA d'un virus,
  - soit par transcription réverse d'un RNA de la cellule ou d'un virus, au moyen d'une transcriptase réverse qui produit un DNA complémentaire (cDNA) et d'une transposase qui l'incorpore au génome de la cellule.



13.2 Protéine rec-A

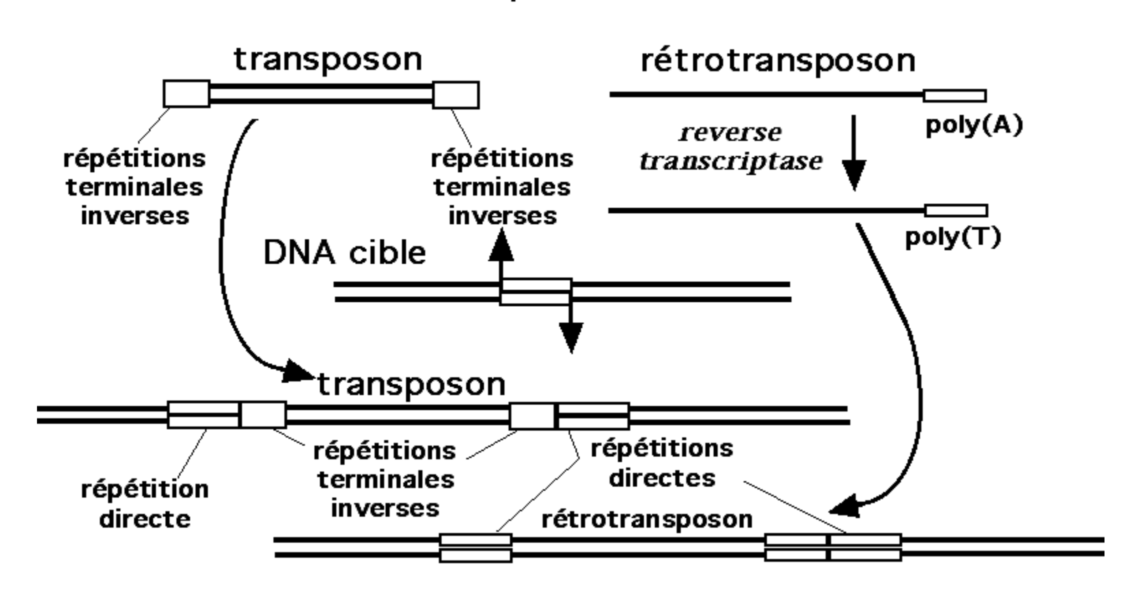


BG 94

- La protéine recA, expression du gène du même nom, joue un rôle essentiel dans le processus de recombinaison homologue, qui permet la réparation de l'ADN chez *Escherichia coli*.
- La protéine recA a pour effet de dissocier les deux brins d'un ADN normal pour permettre l'excision d'une séquence destinée à complèter une brèche d'un ADN dont les deux brins sont lésés irréversiblement.
- Ainsi au cours de la réplication si un des deux ADN produits est anormal et ne contient plus d'information, l'autre ADN, dissocié par recA, détachera un de ses brins pour le recombiner avec l'autre ADN : de sorte qu'il y aura de chaque côté un brin porteur de l'information et que les enzymes de réparation pourront complèter les deux doubles hélices avec la même information.
- Le gène recA est réprimé par la protéine LexA, mais dès que le taux d'ADN simple brin augmente dans la cellule, le catabolisme de cette protéine LexA provoque l'expression de recA pour permettre la réparation des fragments d'ADN lésés (système SOS).



13.3 Transposition (mécanisme)

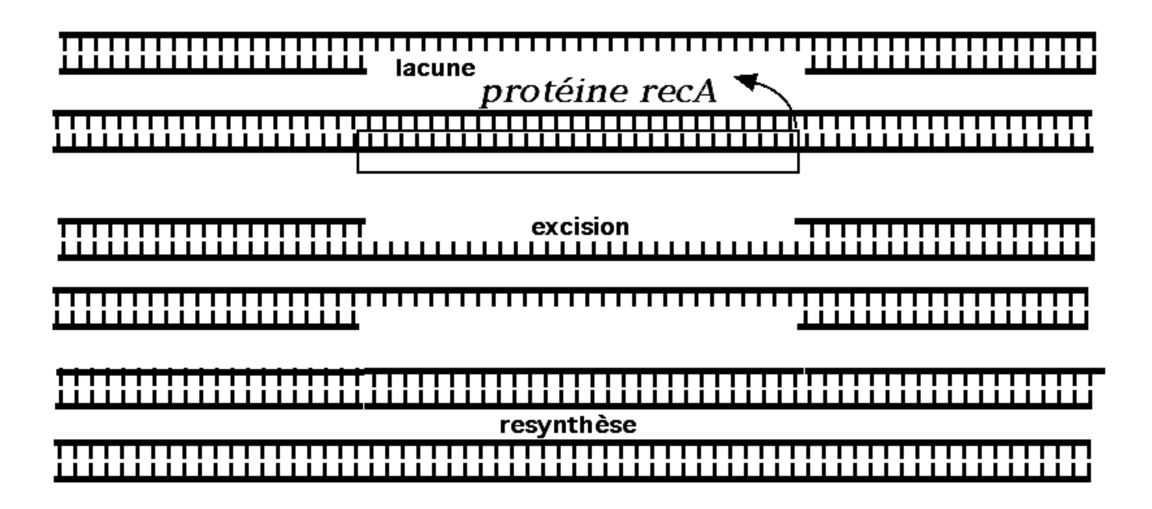


BG 94/1

- La transposition est l'insertion d'une séquence entière d'ADN, provenant d'un ADN viral ou d'un ARN cytoplamique rétrotranscrit.
- Les transposons des virus sont des fragments d'ADN double brin terminés par des séquences répétées inverses (la même séquence de 5'$\rightarrow$3' d'un côté et de 3'$\rightarrow$5' de l'autre côté). Les rétrotransposons sont obtenus par la transcriptase réverse des rétrovirus à partir d'ARN cytoplamiques : les cDNA obtenus seront transposés dans l'ADN de la cellule hôte.
- Dans les deux cas l'ADN du site de transposition sera ouvert par deux brèches sur les brins de l'ADN décalées de dix à vingt nucléotides, créant ainsi une coupure à brins inégaux. Le transposon va s'insérer entre ces deux extrémités, les vides laissés aux extrémités de l'ADN de l'hôte seront comblés par la DNA polymérase et les brèches restantes liées par la DNA ligase.
- Ainsi, l'ADN au niveau d'un transposon viral présente deux répétitions de même sens en dehors du transposon proprement dit qui est encadré par deux autres répétitions de sens inverse. Au niveau d'un rétrotransposon, on rencontre aussi deux répétitions de même sens en dehors du transposon proprement dit qui est terminé par une partie poly(dA:dT) issue de la rétrotranscription de la queue poly(A) de l'ARN d'origine.



13.4 Recombinaison simple (mitose)

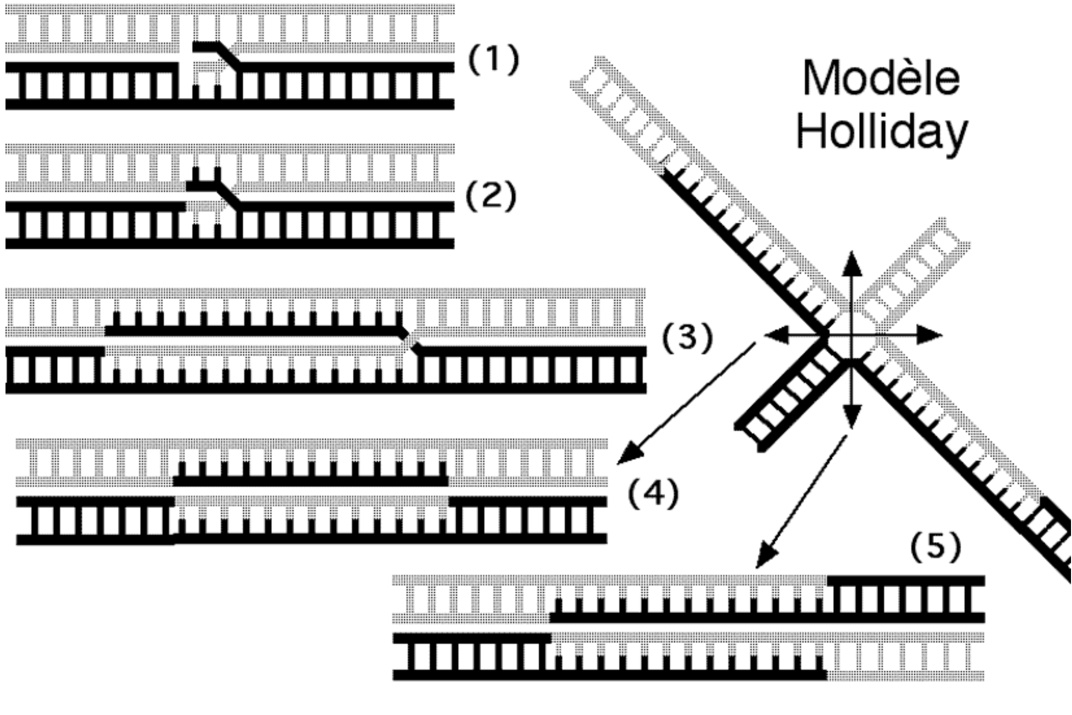


BG 94/2

- Lorsque un brin d'ADN est lésé et qu'il est impossible pour la DNA polymérase de synthétiser l'autre brin, il persiste une lacune. L'existence d'un autre chromosome porteur de la même information génétique, est une source d'information pour permettre la reconstitution du gène altéré.
- Une protéine de recombinaison (analogue de recA) va permettre l'excision d'un des brins du gène sain et la recombinaison du fragment excisé avec le gène présentant une lacune. En même temps le brin lésé sera détruit par les nucléases. A ce stade il y aura deux gènes porteurs d'une information complète sur un seul de leurs brins d'ADN.
- Les enzymes de réparation vont alors reconstruire les brins opposés et refermer les brèches, ce qui conduit à une situation normale.
- Ce type de réparation n'est parfait que pour autant que le gène lésé soit le même sur les deux chromosomes (homozygotie). En cas d'hétérozygotie, le fait de reconstituer les deux gènes à partir du même allèle peut conduire à une « perte d'hétérozygotie », quelquefois source de pathologies nouvelles pour des affections récessives donc inapparentes à l'état hétérozygote.



13.5 Modèle Holliday

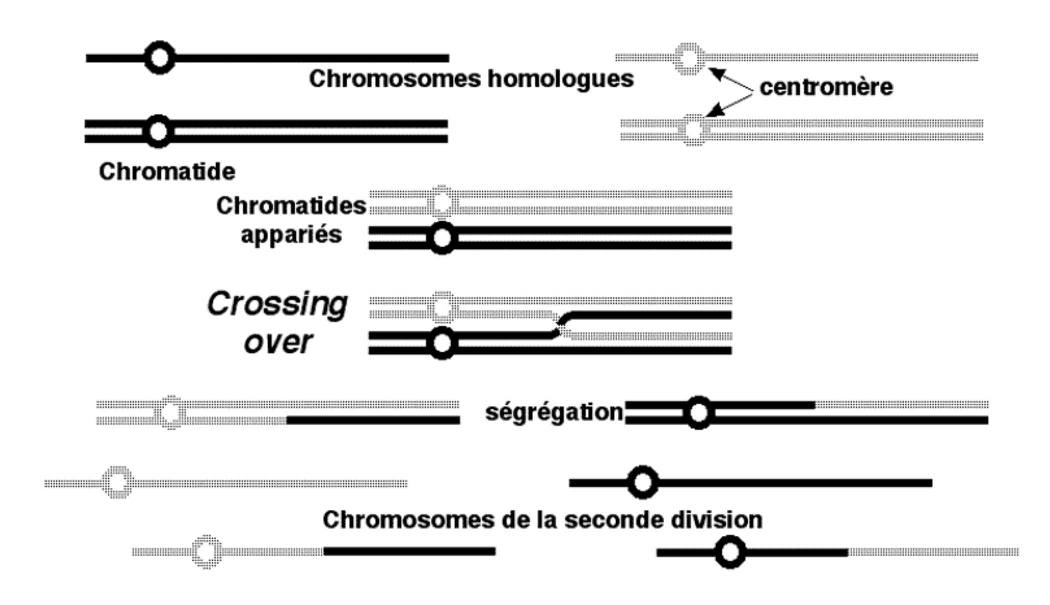


BG 94/3

- L'ouverture d'une brèche dans un brin de DNA et le déroulement de la double hélice permet à des fragments de DNA simple brin de s'hybrider de façon anormale.
- L'existence de séquences homologues sur l'autre chromosome de la même paire peut conduire ce brin libre à s'hybrider avec la séquence complémentaire du chromosome homologue (1).
- Dans ce cas les brèches au bout des deux brins échangés peuvent être refermées par la DNA ligase (2).
- Une telle structure avec entrecroisement de brins de deux DNA homologues est appelée structure Holliday (du nom de son découvreur en 1964). Cette structure est mobile et l'échange peut se propager par glissement de la structure le long des deux hélices en allant dans le même sens (3).
- On peut représenter la structure Holliday en séparant les hélices en forme de croix. Cela permet de voir qu'il y a deux façons de séparer les deux hélices : soit en coupant horizontalement (flèche rouge) et on obtient alors un échange de fragment d'ADN entre les deux chromosomes (4), soit en coupant verticalement (flèche jaune) et on aboutit alors à un échange complet des ADN des deux chromosomes au delà du point de croisement (5).



13.6 Crossing-over



BG 94/4

- Les échanges entre chromosomes homologues peuvent conduire à des échanges de fragments ou à des échanges de brins entiers de DNA.
- Au cours de la méiose en particulier, de tels échanges se produisent fréquemment entre les chromosomes d'origine paternelle ou maternelle de telle sorte que les gènes des chromosomes des cellules filles (gamètes) sont une recombinaison hétérogène de fragments des deux origines.
- Le résultat implique un mélange de caractères héréditaires des deux parents qui constituent le patrimoine du nouvel individu, bien que l'homologie des échanges garantisse le maintien de chaque gène sur chaque locus avec un allèle (homozygote) ou deux (hétérozygote).



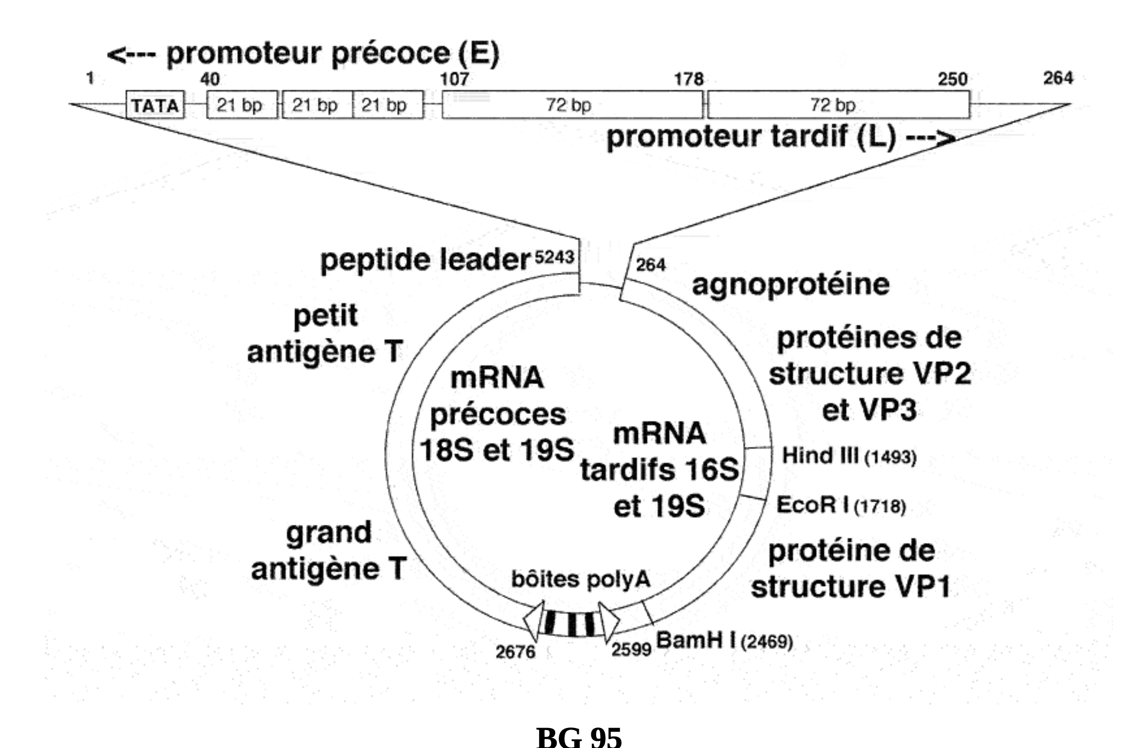

faculté de médecine
Pierre et Marie CURIE

Chapitre 14

Construction de vecteurs



14.1 Carte du virus SV40

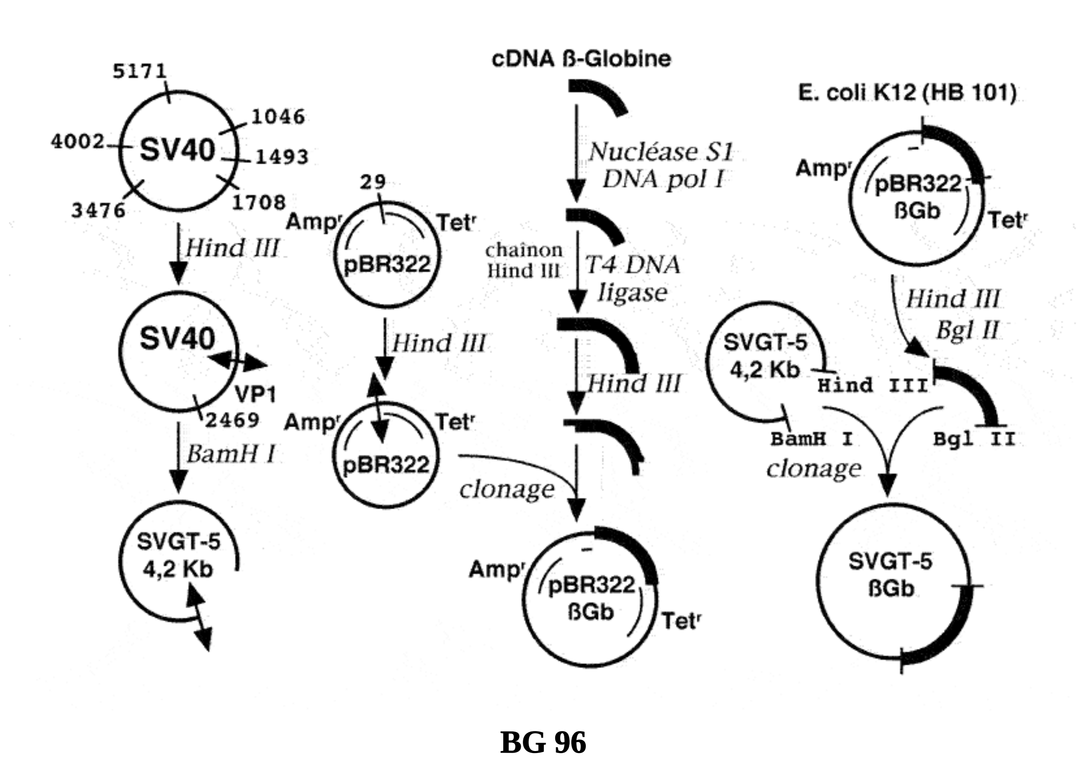


BG 95

- Le virus SV 40 est répliqué très activement par certaines cellules eucaryotes. Au bout de 3 ou 4 jours la cellule contient 10000 copies de cet ADN et meurt. Les vecteurs dérivés de ce virus permettent donc une expression abondante mais transitoire des gènes qu'ils transportent.
- Le génome du virus est un ADN double brin circulaire de 5243 paires de bases (pb).
- L'ADN viral est transcrit en un messager précoce qui par épissage alternatif donne deux antigènes T qui sont des facteurs trans-activateurs de la réplication du virus. Après la réplication du virus un autre promoteur conduit à la transcription des messagers tardifs qui codent pour 3 protéines de la capside virale.



14.2 Construction d'un vecteur



BG 96

- Le clonage de la β-globine dans un vecteur eucaryote a été décrit par Mulligan, Howard et Berg (Nature 1979; 277: 108-114).
- Le DNA du virus SV 40 a été digéré partiellement par Hind III, de façon à ne produire qu'une brèche à la position 1708, qui correspond au site d'initiation de la traduction de la protéine virale VP1. Puis le produit a été digéré à nouveau par BamH I qui n'a qu'un seul site en 2469 proche du site poly(A) du messager de cette protéine. Le produit est un vecteur d'expression nommé SVGT-5 de 4,2 Kb.
- Le gène de la β-globine du lapin a été isolé. On a fait des bouts francs grâce à la nucléase S1 et à la DNA polymérase de E.coli. Sur ces bouts francs on a lié avec la DNA ligase de T4, un chainon d'ADN double brin, à bouts francs, de séquence CCAAGCTTGG, comportant un site de restriction pour Hind III (chaînon Hind III). L'ADN ainsi construit a été digéré par cette enzyme de restriction pour produire des bouts collants Hind III.
- Le plasmide pBR322 a été également digéré par Hind III en un site unique et le clonage de l'ADN de globine préparé ci-dessus a été fait dans cette ouverture Hind III. Le plasmide obtenu a servi à transfecter des colonies de *E. coli* K12 (HB 101) qu'on a fait pousser en présence d'ampicilline. On a sélectionné ensuite des colonies qui n'étaient plus résistantes à la tétracycline : en effet le plasmide recombiné pBR322-βGb confère la résistance à l'ampicilline mais ne confère plus la résistance à la tétracycline.
- On récupère l'insert amplifié en digérant les plasmides par Hind III et par Bgl II, cette dernière enzyme possédant un site unique à la fin de l'ADN de la β-globine. Enfin on a pu recom-



biner cet ADN de globine avec le vecteur SVGT-5 parce que les bouts collants des enzymes BamH I (g/gatcc) et Bgl II (a/gatct) ont la même séquence.



14.3 Vecteurs hybrides

Vecteurs hybrides

- **Phagémides**
  - Phage filamenteux + plasmide
  - exemples : pBluescript II, vecteur T
  - inclus dans un phage λ = λ ZAP II
- **Cosmides**
  - plasmide + phage λ (grands inserts)
- **Vecteurs navettes (hôtes eucaryotes et procaryotes)**
- **Chromosomes artificiels de levure (YAC)**

BG 97

- La recombinaison de l'ADN permet de construire des vecteurs artificiels qui associent les propriétés des phages et celles d'autres vecteurs.
- Les phagémides sont des constructions associant l'ADN d'un phage et celui d'un plasmide bactérien.
- Pour permettre l'insertion de fragments d'ADN de grande longueur, on utilise des cosmides, construits à partir de phage λ et de plasmides ou bien , dans le cas des cellules eucaryotes, des YAC (*Yeast Artificial Chromosome* = chromosome artificiel de levure).
- Lorsque la manipulation comporte le passage d'une amplification chez un procaryote suivi d'une expression chez un eucaryote, on peut utiliser des vecteurs-navettes capables de transporter l'insert d'une cellule à l'autre.




faculté de médecine
Pierre et Marie CURIE

Chapitre 15

Transgénèse



15.1 Vecteurs de thérapie

Vecteurs de thérapie

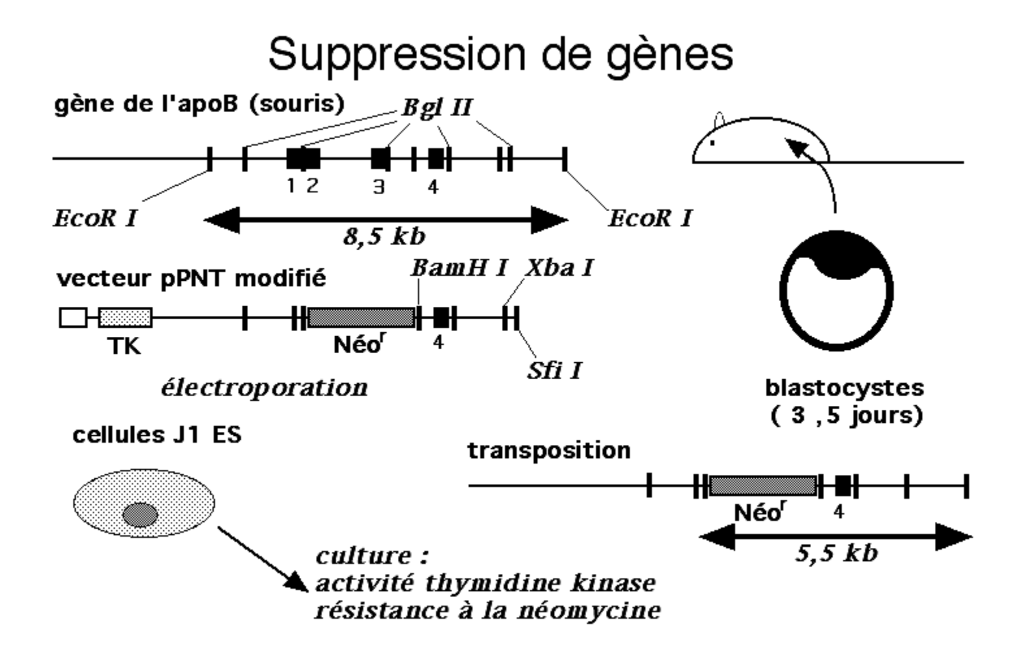
- **Rétrovirus**
- **Adénovirus**
- **Herpés**
- **Liposomes**

BG 98

- La génétique moléculaire, lorsqu'elle étudie la pathologie humaine, a bien sûr pour objectif de pallier les conséquences des lésions génétiques héréditaires et pathogènes : c'est le domaine de la thérapie génique.
- Lorsqu'un gène n'est pas exprimé dans les cellules d'un individu, on a pour objectif de transfecter ses cellules avec un vecteur qui apportera le gène normal dans un environnement permettant son expression et la régulation de cette expression.
- Les vecteurs utilisables pour transfecter des fragments d'ADN, *in vivo* chez un malade, sont de plusieurs catégories :
  - — des rétrovirus à ARN, pourvus d'une transcriptase réverse capable de transposer le cDNA de leur génome dans les chromosomes humains ;
  - — des adénovirus à ADN, qui font exprimer préférentiellement leur génome dans les cellules infectées ;
  - — le virus de l'herpès, maladie bénigne due à un virus qui peut servir de vecteur ;
  - — des liposomes, sortes de sphères limitées par une membrane lipidique artificielle, qui sont captés spécifiquement par certaines cellules, grâce à des récepteurs de particules lipidiques.



15.2 Souris *knock out*



BG 99

- L'étude des fonctions engendrées par l'expression d'un gène peut se faire en supprimant le gène du patrimoine génétique d'un animal d'expérience. Une souris ainsi privée du gène de l'apolipoprotéine B est appelée souris knock-out pour le gène apoB (Li-Shin Huang : J.Clin.Invest. 1995; 96: 2152-61).
- L'ADN visé pour supprimer l'expression de l'apoB est le début du gène de cette protéine, en particulier les quatre premiers exons qui sont compris dans un fragment EcoR I de 8,5 Kb.
- On prépare un vecteur (pPNT modifié) qui comprend un promoteur, un gène reporter (thymidine kinase = TK) et un gène de résistance à la néomycine (Néor) et une partie du gène de l'apoB (jusqu'à l'exon 4). Ce vecteur est amplifié dans des cellules J1 ES, testée pour la présence de l'activité thymidine kinase et la résistance à la néomycine.
- Des œufs fécondés de souris, au stade blastocyste (après 3,5 jours de développement) sont transfectés par le vecteur. L'ADN du vecteur partiellement homologue du gène de l'apoB, va se transposer à la place de ce gène sur une longueur de 5,5 Kb, pour donner un gène recombinant dans lequel le gène de résistance à la néomycine a remplacé les premiers exons du gène de l'apoB. Les œufs sont réimplantés dans un utérus de la mère et la gestation se poursuit.
- Le gène apoB modifié ne sera plus exprimé et il en résultera une absence totale de lipoprotéines de basse densité chez les souris qui auront subi cette transposition.

